



THE ECONOMICS OF **Imperfect Labor Markets**

Tito Boeri and Jan van Ours

The Economics of Imperfect Labor Markets

This page intentionally left blank

The Economics of Imperfect Labor Markets

Tito Boeri and Jan van Ours

Princeton University Press
Princeton and Oxford

Copyright © 2008 by Princeton University Press

Published by Princeton University Press

41 William Street, Princeton, New Jersey 08540

In the United Kingdom: Princeton University Press

6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

All Rights Reserved

Library of Congress Cataloging-in-Publication Data

Boeri, Tito.

The economics of imperfect labor markets / Tito Boeri and Jan van Ours.

p. cm.

Includes bibliographical references (p.) and index.

ISBN: 978-0-691-12449-0 (alk. paper) 978-0-691-13735-3 (pbk. : alk. paper) 1. Labor market. I. Ours, J. C. van. II. Title.

HD5706 .B648 2008

331.12—dc22

2008062104

British Library Cataloging-in-Publication Data is available

This book has been composed in Times Roman

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Contents

Preface	xi
Symbols and Acronyms	xvii
1 Overview	1
1.1 A Few Key Definitions	2
1.2 A Competitive Labor Market	4
1.2.1 Labor Supply and the Reservation Wage without Hours Restrictions	4
1.2.2 Aggregate Labor Supply	7
1.2.3 Labor Demand and Wage Determination	9
1.2.4 Equilibrium	11
1.3 Labor Market Institutions	14
1.3.1 Acting on Prices	14
1.3.2 Acting on Quantities	15
1.3.3 Institutional Interactions	17
1.3.4 Why Do Labor Market Institutions Exist?	18
1.3.5 Product Market Competition and Institutional Reforms	21
1.4 Technical Annex: A Simple Static Framework	24
1.4.1 A Competitive Labor Market	24
1.4.2 Labor Market Institutions	25
1.4.3 The Wedge	26
1.4.4 Product Market Competition and the Employment Bias of Institutions	27
2 Minimum Wages	29
2.1 Cross-Country Comparisons	30
2.2 Theory	33
2.2.1 A Competitive Labor Market	33
2.2.2 A Noncompetitive Labor Market	34
2.3 Empirical Evidence	38
2.3.1 Studies Based on Firm-Level Data	38
2.3.2 Studies Based on Natural Experiments	39
2.3.3 Studies Based on Workers' Histories	43
2.4 Policy Issues	44
2.4.1 Should the Minimum Wage Be Reduced or Increased?	44
2.4.2 Is the Minimum Wage Effective in Reducing Earnings Inequality and Poverty?	45
2.5 Why Does a Minimum Wage Exist?	46

2.6	Suggestions for Further Reading	47
2.7	Review Questions	47
2.8	Technical Annex: Fine-Tuning of the Minimum Wage	48
3	Unions and Collective Bargaining	51
3.1	Measures and Cross-Country Comparisons	52
3.1.1	Union Density	52
3.1.2	Coverage and Excess Coverage	54
3.1.3	Coordination	56
3.2	Theory	56
3.2.1	Collective Bargaining	57
3.2.2	Endogenous Membership	61
3.3	Empirical Evidence	66
3.3.1	Effects of Unions on Wages	67
3.3.2	Bargaining Coordination, Union Density, and Unemployment	71
3.4	Policy Issues	72
3.4.1	Do Unions Increase Efficiency?	72
3.4.2	Should Collective Bargaining Be Decentralized?	74
3.5	Why Do Unions Exist?	75
3.6	Suggestions for Further Reading	76
3.7	Review Questions	76
3.8	Technical Annex: How Strong Should Unions Be in Order to Be Efficient?	77
4	Payroll Taxes	81
4.1	Measures and Cross-Country Comparisons	82
4.2	Theory	86
4.3	Empirical Evidence	90
4.4	Policy Issues	92
4.4.1	Negative Income Taxes or In-Work Benefits?	92
4.4.2	Tax Credits or Wage Subsidies?	96
4.5	Why Do Payroll Taxes Exist?	97
4.6	Suggestions for Further Reading	98
4.7	Review Questions	98
4.8	Technical Annex: Taxes and Benefits in a Competitive Labor Market	98
5	Regulation of Working Hours	101
5.1	Measures and Cross-Country Comparisons	101
5.2	Theory	102
5.2.1	Regulating Weekly Working Hours	104
5.2.2	Part-Time Work	107
5.3	Empirical Evidence	109
5.3.1	Working Hours	109
5.3.2	Part-Time Work	113
5.4	Policy Issues	116
5.4.1	Should Governments Regulate Working Hours?	116
5.4.2	Should Governments Stimulate Part-Time Labor?	117
5.5	Why Does Regulation of Working Hours Exist?	117
5.6	Suggestions for Further Reading	118
5.7	Review Questions	118
5.8	Technical Annex: Reduction of Standard Working Hours	119

6 Retirement Programs	121
6.1 Measures and Cross-Country Comparisons	122
6.2 Theory	124
6.3 Empirical Evidence	125
6.3.1 Age and Employment	125
6.3.2 Age and Productivity	132
6.4 Policy Issues	134
6.4.1 Should the Mandatory Retirement Age Be Increased?	134
6.4.2 Should Early Retirement Programs Be Phased Out?	135
6.5 Why Do Early Retirement Programs Exist?	136
6.6 Suggestions for Further Reading	136
6.7 Review Questions	136
6.8 Technical Annex: Optimal Retirement Age	137
7 Family Policies	139
7.1 Measures and Cross-Country Comparisons	140
7.2 Theory	142
7.2.1 Child Care Facilities	142
7.2.2 Parental Leave	142
7.3 Empirical Evidence	144
7.4 Policy Issues	150
7.4.1 Can Work and Family Life Be Balanced?	150
7.4.2 Is There a Trade-off between Fertility and Employment?	151
7.5 Why Do Family Policies Exist?	153
7.6 Suggestions for Further Reading	153
7.7 Review Questions	154
7.8 Technical Annex: Child Care Facilities and Hours of Work	154
8 Education and Training	157
8.1 Measures and Cross-Country Comparisons	158
8.2 Theory	160
8.2.1 Schooling	160
8.2.2 Training	163
8.3 Empirical Evidence	165
8.3.1 Returns to Schooling	165
8.3.2 On-the-Job Training	169
8.4 Policy Issues	170
8.4.1 Should Governments Subsidize In-Company Training?	170
8.4.2 Should There Be a Compulsory Schooling Age?	171
8.5 Why Do Governments Provide Education and Training?	171
8.6 Suggestions for Further Reading	172
8.7 Review Questions	172
8.8 Technical Annex: Schooling and Training	172
8.8.1 Optimal Years of Schooling	172
8.8.2 Who Pays for General Training?	174
9 Migration Policies	175
9.1 Measures and Cross-Country Comparisons	176
9.2 Theory	178

9.2.1	A Competitive Labor Market	179
9.2.2	An Economy with Wage Rigidities	180
9.2.3	Wage Rigidities and Unemployment Benefits	181
9.2.4	What Drives Migration Decisions?	182
9.2.5	Effects on Income Distribution of Skill-Biased Migration	183
9.3	Empirical Evidence	184
9.3.1	The Labor Market Performance of Migrants	189
9.3.2	Fiscal Effects	190
9.4	Policy Issues	191
9.4.1	Closing the Welfare Door?	192
9.4.2	Adopting a Points System?	193
9.5	Why Do Migration Policies Exist?	194
9.6	Suggestions for Further Reading	195
9.7	Review Questions	195
9.8	Technical Annex: Net Gains from Migration and the Option Value of Waiting	195
10	Employment Protection Legislation	199
10.1	Measures and Cross-Country Comparisons	200
10.2	Theory	204
10.2.1	A Neutrality Result	204
10.2.2	Removing Risk Neutrality	206
10.2.3	EPL with Rigid Wages	206
10.2.4	EPL as a Tax	209
10.2.5	Two-Tier Regimes	210
10.3	Empirical Evidence	211
10.3.1	Cross-Country Analyses	211
10.3.2	Within-Country Studies	213
10.3.3	Endogeneity of EPL	214
10.4	Policy Issues	215
10.4.1	How Much Protection Should EPL Provide?	215
10.4.2	Whom Should EPL Be Protecting?	218
10.5	Why Does Employment Protection Legislation Exist?	220
10.6	Suggestions for Further Reading	220
10.7	Review Questions	221
10.8	Technical Annex: EPL in a Search-Matching Model	221
10.8.1	A Simple Dynamic Framework	221
10.8.2	Present Discounted and Flow Values	221
10.8.3	Behavior of Workers	222
10.8.4	Matching	222
10.8.5	Firms	223
10.8.6	Wage Bargaining	223
10.8.7	Introducing EPL	223
10.8.8	Baseline Parameters	224
11	Unemployment Benefits	225
11.1	Measures and Cross-Country Comparisons	226
11.2	Theory	230
11.2.1	A Competitive Labor Market	230
11.2.2	An Imperfect Labor Market	234

11.3	Empirical Evidence	238
11.3.1	Effects on Reservation Wage	238
11.3.2	Effects on Unemployment Duration: Aggregate Data	239
11.3.3	Effects on Unemployment Duration: Microeconomic Data	240
11.3.4	Effects of Unemployment on the Generosity of UBs	243
11.4	Policy Issues	244
11.4.1	Why and When Should UBs Be Publicly Provided?	244
11.4.2	What Is the Optimal Structure of UBs?	246
11.5	Why Do Unemployment Benefits Exist?	247
11.6	Suggestions for Further Reading	248
11.7	Review Questions	248
11.8	Technical Annex: UBs in a Search-Matching Model	249
11.8.1	Effects of UBs in a Two-Sided Job Search Model	249
11.8.2	Optimal Job Search Intensity	249
11.8.3	Vacancies and the Matching Function	250
11.8.4	Wage Bargaining	251
11.8.5	Labor Market Participation	251
11.8.6	Balanced Budget	253
11.8.7	Numerical Simulations	253
12	Active Labor Market Policies	255
12.1	Measures and Cross-Country Comparisons	256
12.2	Theory	258
12.3	Empirical Evidence	263
12.3.1	Experimental Studies	264
12.3.2	Nonexperimental Studies	268
12.4	Policy Issues	270
12.4.1	Do We Need Public Employment Services?	271
12.4.2	Do We Need Activation Policies?	271
12.5	Why Do Active Labor Market Policies Exist?	272
12.6	Suggestions for Further Reading	272
12.7	Review Questions	273
12.8	Technical Annex: Activating Unemployed Workers	273
13	Institutional Interactions	277
13.1	Taxes and Unions	278
13.2	Employment Protection and Unemployment Benefits	280
13.3	Unemployment Benefits and Active Labor Market Policies	282
13.4	Employment Protection and Unions	284
13.5	Taxes and Unemployment Benefits	285
13.6	Education and Retirement Programs	287
13.7	What Future for These Institutions?	288
	References	291
	Index	313

This page intentionally left blank

Preface

Is it good for employment to increase the progressivity of taxation? Does it make sense to contrast “active” and “passive” labor market policies? Who gains and who loses from employment protection legislation? Why are minimum wages generally diversified by age? Is it better to have decentralized or centralized bargaining systems in monetary unions? Should migrants have access to welfare benefits? Should governments regulate working hours?

Current labor economics textbooks do not address these relevant policy issues. In spite of significant progress in analyzing the costs and benefits of labor market institutions, these books have a setup that relegates institutions to the last paragraph of chapters or to a final institutional chapter. Typically a book begins by characterizing labor supply (including human capital theory), labor demand, and the functioning of the labor market and subsequently addresses such topics as wage formation and unions, compensating wage differentials, discrimination, and unemployment. There is little information concerning labor market institutions and labor market policy. Usually labor market policies are mentioned only every now and then, and labor market institutions are often not treated in a systematic way. When attention is given to labor market institutions, reference is generally made to the U.S. institutional landscape and to competitive labor markets in which, by definition, any type of policy measure is distortional.

The novelty of this book is that from the start the focus is on labor market institutions operating in *imperfect* labor markets, that is, markets that depart from perfect competition. Imperfect markets are characterized by the presence of many labor market *institutions*, that is, systems of laws and programs that shape the behavior of individual workers and employers. Institutions result from a political process aimed at (1) increasing economic efficiency and (2) achieving some redistributive goal. Efficiency is achieved by remedying market imperfections, such as excessive monopsonistic power, informational asymmetries that give rise to moral hazard and adverse selection problems, and externalities associated with social customs or the job-matching process, as well as transaction costs and frictions that restrict the size of markets. Redistribution provides a rationale for these institutions even when there are no market imperfections. In imperfect markets redistribution sometimes can be achieved while pursuing efficiency, as in the case of institutions, such as

minimum wages or employment subsidies, that counteract excessive monopsonistic power. In most cases, however, the traditional trade-off between efficiency and equity arises. Actually, the redistribution operated by these institutions may well not promote a more egalitarian society or represent the interests of the median voter. There are frequent policy failures in the design of labor market institutions that give disproportionate representation to some pressure group pursuing very specific interests.

This book also takes into account that institutions rarely operate in isolation. Hence, from a positive standpoint, the effects of each institution on the labor market are investigated by considering not only its direct effects on employment, unemployment, and wages, but also its indirect effects, mediated by the presence of other institutions. For example, a change in the generosity of the unemployment benefit system affects unemployment directly by reducing search intensity and increasing the reservation wage of jobseekers and indirectly by increasing the bargaining power of unions and the level of the efficiency wage. This, incidentally, provides a third rationale for the presence of some institutions: they are created to counteract or complement the effects of other institutions. Policy failures may arise also in this context because the institutions that are responsible for the distortions are rarely reformed. Often the political process creates chains of distortions whereby institutions are used to compensate for undesirable effects of other institutions.

Much attention is placed on precisely defining institutions and measuring them along their relevant dimensions (e.g., eligibility for unemployment benefits, level of the benefits, and maximum duration for which they can be provided) because we believe that accuracy in describing the way in which the institution operates and the goals it pursues is essential in characterizing its effects on the labor market. Statistical information on the evolution over time of these institutions is also provided, whenever possible, for all OECD countries. Contrary to common wisdom, there has indeed been considerable variation over time in these institutions. Policy enforcement issues are also discussed.

Position of This Book in the Literature

As mentioned earlier, available labor economics textbooks devote one or a few chapters to institutions. Borjas (2006), Kaufman and Hotchkiss (2006), Ehrenberg and Smith (2006), McConnel et al. (2008) are just a few examples. In these chapters the key distortions associated with the presence of labor market institutions are only briefly touched upon, and there is no attempt to examine how the institutions operate when labor markets are not competitive.

A few specialized books analyze labor market institutions. Because they are not conceived as textbooks, they do not provide systematic coverage of the different institutions and do not offer an integrated framework. A partial exception

is Layard, Nickell, and Jackman (2005), which, however, is confined to the analysis of unemployment. Manning (2003) is another partial exception: it offers a clear overview of monopsony in all its dimensions. However, except for unions and minimum wages, it does not deal with institutions in much detail. Finally, Cahuc and Zylberberg (2004) offer wider and in-depth coverage of labor market institutions, but they fail to provide a unifying framework, and their book is, in any event, intended for graduate students.

There is a huge literature on the effects of institutions on labor market performance. This book does not aim at accounting for all works that have been written on the subject. We concentrate on those findings and issues that have, in our view, more relevance in real-world labor markets.

We do not cover a number of topics that are frequently included in labor economics textbooks, such as discrimination and performance-related pay. Although we discuss why institutions exist, we do not offer full coverage of the political economy of labor market institutions, which is addressed by other specialized books (Saint-Paul 2000; Persson and Tabellini 2000).

Audience

This book aims at reaching a composite readership, including undergraduate students taking courses in labor economics (a compulsory requirement for many BAs), as well as graduate students specializing in this field. Professional economists in international organizations and government agencies are another potential target.

This book can be taught at the undergraduate level in programs that specialize in economics and business. It can also be used in business and political science schools where human resource-oriented courses are taught by economists. The description of the institutions and the technical annexes can also be useful references in graduate labor economics courses, which can also benefit from the datasets and applications provided on our webpage (see link at <http://press.princeton.edu/titles/8771.html>) and referred to in the book. Finally, scholars in the field may find the book a useful reference text for their personal library.

Prerequisites and Technical Level

In light of the audience in mind, the technical level required by the book is modest. Ideally, readers should have taken an introductory course in microeconomics, a semester of calculus, and an introductory course in statistics. In practice it will be possible to read the book even when these prerequisites are partially or totally not fulfilled. The viability of the latter option rests on the various numerical examples presented in the book and on the graphical treatment of some key results. In this

simpler treatment all the main arguments are presented and the main results are outlined, even though they lack the rigor and the generality that the use of calculus allows. Boxes and, above all, technical annexes provide these features.

Plan and Guidelines for Instructors

This book consists of 13 chapters, including an overview chapter and a final chapter on institutional interactions. Each chapter except the first and the last focuses on a different institution. The overview chapter sets out the unifying line of reasoning and the structure of the book and dwells on the relationships between labor markets and product markets, notably the effects of increased product market competition on labor market institutions. Then the institutions are discussed separately in 11 chapters. Each chapter can be dealt with separately. The theoretical framework is formally presented in technical annexes. There are some repetitions in the annexes, but the advantage is that each chapter is self-contained (except the overview and the final chapter). It can be used as teaching material without necessarily going through the previous chapters.

This organization of the book allows the instructor to choose any particular sequence of institutions for a smaller course. For instance, a basic course (not requiring dynamic frameworks) on *price-based institutions* could cover chapters 2, 3 and 4; a basic course on *quantity-based institutions* could cover chapters 5, 6 and 7. More advanced courses, requiring some dynamic modeling, could focus on chapters 8 through 12. Smaller courses can also be organized by topics, for example a course on *flexicurity* (chapters 10, 11, 12, and 13), a course on *wage-compressing institutions* (chapters 2 and 3), or a course on *human capital investment* (chapters 8 and 9).

In all chapters except chapter 7, where we concentrate on family labor supply decisions, the analytical unit is the individual. We also treat labor supply decisions along extensive margins (participation) in all chapters except chapter 4 and chapter 5, where we consider also adjustment along the intensive margin (hours of work).

Each of the 11 chapters that concerns a separate institution is set up in the same way:

1. Definition of the institution and the way in which it is enforced
2. Measurement issues and stylized facts about the institution (cross-country variation and time-series evolution)
3. Theories (the plural is used because for each institution several theories may apply)
4. Empirical evidence (macro and micro evidence)
5. Policy issues (relevant trade-offs and design features)

6. Overall assessment and rationale for the presence of the institution
7. Suggestions for further reading
8. Review questions
9. Technical annex

Distinguishing Features

Two distinguishing features of our approach are (1) a thorough discussion of measurement issues that tries, wherever possible, to complement institutional indicators with information on the actual enforcement of these norms and (2) an attempt to highlight the rationale behind each institution, its efficiency and distributional properties, and those who benefit and those who lose from its presence.

In the review of the empirical evidence, one or two studies are selected (often applying a difference-in-differences methodology to isolate the effects of a particular institution) and discussed in greater detail in a box. We also provide on our webpage (see link at <http://press.princeton.edu/titles/8771.html>) the links to the original study and the data and programs in Stata that can be used to replicate results of the study, test their robustness, and potentially extend them in new directions. This allows students to deepen their knowledge of institutional data and measure and to better learn the econometric techniques that can be best used to analyze the effects of these institutions, nicely complementing their training in microeconometrics.

Suggestions for further reading aim at complementing this with a few seminal publications. Detailed references are provided at the end of this book.

This book has been tested during various courses that the authors gave at their universities and at international organizations. The lecture notes of these courses became the initial outlines of the various chapters, and professional economists and students attending our courses provided an excellent sounding board.

Acknowledgments

Most of the coordination work on this book was done while we were hosted by the Institute for the Study of Labor (IZA) in Bonn. We are very grateful to Klaus Zimmermann, the director of IZA, and Ulriche Maurer for their warm hospitality. The Netherlands Institute for Advanced Studies (NIAS) offered a very pleasant environment for the final editing of this book. Bocconi University also supported our work with a grant for improvement of the quality of teaching. Battista Severgnini provided unflagging and skillful research assistance and helped us also in the preparation of the webpage. Andrew Glyn kindly allowed us to have access to the long series of institutions he skillfully collected in his work.

We are particularly indebted to the colleagues who agreed to supply the datasets that are now available on the webpage and allow students to replicate the results reviewed in the book. In particular, we would like to warmly thank Orley Ashenfelter, Andrea Bassanini, Richard Blundell, Espen Bratberg, David Card, Romain Duval, Tor-Helge Holmås, Jennifer Hunt, Dean R. Hyslop, Juan Jimeno, Alan Krueger, Rafael Lalive, Ghazala Naz, David Neumark, Thomas Piketty, Cecilia Rouse, Øystein Thøgersen, Milan Vodopivec, and William Wascher.

While we were finalizing the book, we benefited from insightful comments of anonymous referees of individual chapters and three external readers of the entire manuscript. It was also very useful for us to have the opportunity to give short courses at the European Commission and the International Labor Organization in order to improve cross-country comparisons of institutions. Roberta Marcaletti skillfully assisted us in the final editing of the book. We are also grateful to Richard Baggaley, economics editor at Princeton University Press, for having supported this project from the start without putting too much pressure on us.

Symbols and Acronyms

Symbols

a	per worker lump-sum tax or employment subsidy
a_e	valuation of employer taxes by employees
a_w	discounting of employee taxes by employees
A	index of productivity of matching function or production function
b	unemployment benefit
b^w	replacement rate (of unemployment benefits)
B	bond
c	consumption
C	total labor costs of the firm
c_c	costs of childcare
c_f	firing cost (deadweight)
C_0	migration costs
c_s	costs of schooling
c_v	cost of vacancy
c_τ	cost of general training
e	effort and employment rate
E	expectation
e_g	number of government officials monitoring the unemployed
F	fixed labor costs
F_c	fixed costs of childcare
f_l	marginal product of labor
$G(.)$	cumulative density function of reservation wages
F_{in}^u	inflow into unemployment
h	hours of work
$\bar{h}$	standard working week
h_{ft}	hours of full-time job

h_{pt}	hours of part-time job
i	market interest rate
J_e	value of a job for the firm
J_v	value of a vacancy
L	employment
L^d	labor demand
L_m	equilibrium employment under monopsony
L^s	labor supply
l	leisure
l_0	endowment of time of the individual
m	nonlabor income or flow of matches
M	migrants
N	total population of working age
OV	option value
σ	penalty
q	rate at which shirking workers are detected and fired
r	age of retirement or rate of return to schooling
s	search intensity or years of schooling
t	tax rate
T	total tax or expected length of life
t_e	tax nominally charged to the employee
t_f	tax nominally charged to the firm
T^e	expected length of the working life
u	unemployment rate
U	number of unemployed individuals
u_s	unemployed with benefit sanction
u_u	unemployed without benefit sanction
v	vacancy rate
V	number of vacancies
V_e	value of being employed
V_e^N	value of a job for a nonshirking worker
V_e^S	value of a job for a shirking worker
V_u	value of unemployment
w	hourly wage
$\underline{w}$	minimum wage
w^b	wage under bilateral bargaining

w^e	efficiency wage
w^g	wage in government sector
w^m	wage under monopsony
w^r	reservation wage
w_{ft}^r	reservation wage with full-time hours constraint
w^u	wage under monopoly union
y	value of a job, value of the marginal product of labor
Y	annual earnings
α	parameter of utility function or matching function
β	bargaining power of workers
β_m	union member wage premium
$\gamma(s)$	search costs (s is search intensity)
δ	subjective discount factor
δ^s	job separation rate
ε	inverse labor supply elasticity
η	inverse labor demand elasticity or parameter of matching function
θ	labor market tightness (ratio of vacancies to unemployed workers)
k	parameter of the search cost function
Λ	Lagrangian multiplier
μ	parameter for job-finding rate or wedge
$\bar{\mu}$	exogenous job-finding rate
π	profit
ρ	discount factor (in continuous time)
τ	amount of general training
ϕ	monitoring intensity
ξ	constants of integration
ω	overtime premium

Acronyms

ALMP	active labor market policies
APW	average production worker
BC	Beveridge curve
DB	defined benefit
DC	defined contribution
ECHP	European Community Household Panel
ECVTS	European Continuing Vocational Training Survey

EITC	earned-income tax credit
EPL	employment protection legislation
GDP	gross domestic product
IALS	International Adult Literacy Survey
ISCED	International Standard Classification of Education
ISSP	International Social Survey Programme
MIG	migration policies
NEB	Nonemployment benefit
NIT	negative income tax
NPV	net present value
OECD	Organization for Economic Cooperation and Development
OV	option value
PAYG	pay-as-you-go
PES	public employment service
PPP	purchasing power parity
TR	transfer component of EPL
TWA	temporary work agency
UA	unemployment assistance
UB	unemployment benefit
UI	unemployment insurance
VAT	value added tax
VMP	value of the marginal product of labor
WT	working-time regulations

The Economics of Imperfect Labor Markets

This page intentionally left blank

1

Overview

A large body of academic papers and policy reports examines the effects of labor market institutions on economic performance. This literature was inspired by transatlantic comparisons of employment/unemployment performance and draws mainly on cross-country comparisons. The most influential policy report—the *OECD Jobs Study* commissioned by the Group of Seven (G7) in the early 1990s and completed in 1994—is an attempt to explain the dismal employment/unemployment performance of Europe vis-à-vis the U.S. “jobs miracle.” The key message provided by this report, as well as by many subsequent cross-country studies and policy reports, is that there are in Europe institutional “rigidities” that prevent the labor market from creating as many jobs in the private sector as in the United States. Many academic researchers followed the same route in analyzing various dimensions of the so-called Eurosclerosis, for example, Bean (1994), Alogoskoufis et al. (1995), Snower and de la Dehesas (1996), Nickell (1997), Nickell and Layard (1999), Blanchard and Wolfers (2000), Nickell, Nunziata, and Ochel (2005), and Blanchard (2006). This literature offers a wealth of facts and theoretical results on the relationship between institutions and labor markets. In this book we draw extensively on these findings.

Two problems with this literature, which are acknowledged also in the 2006 reassessment of the *OECD Jobs Study* (OECD 2006c), are worth mentioning at the outset. The first problem is that this literature often fails to explain why these institutions are in place to start with. Institutions are described as something that distorts the work of the market mechanism and prevents the attainment of efficient outcomes. If a government could remove these institutions, it should do so without further ado. It is an offense to the rationality of citizens and of their democratically elected governments that most of these institutions still exist.

The second problem is that this literature tends to have a short memory. The focus is on the recent labor market performance of Europe vis-à-vis the United States. But the institutions that bear the brunt of blame for the poor employment performance of Europe were there also 30 to 40 years ago when the fate of labor markets seemed to be the other way around. Consider figure 1.1: it was only in the mid-1980s, after two oil shocks (the first two vertical lines in the figure), that unemployment in Europe started rising above U.S. levels, and it took another

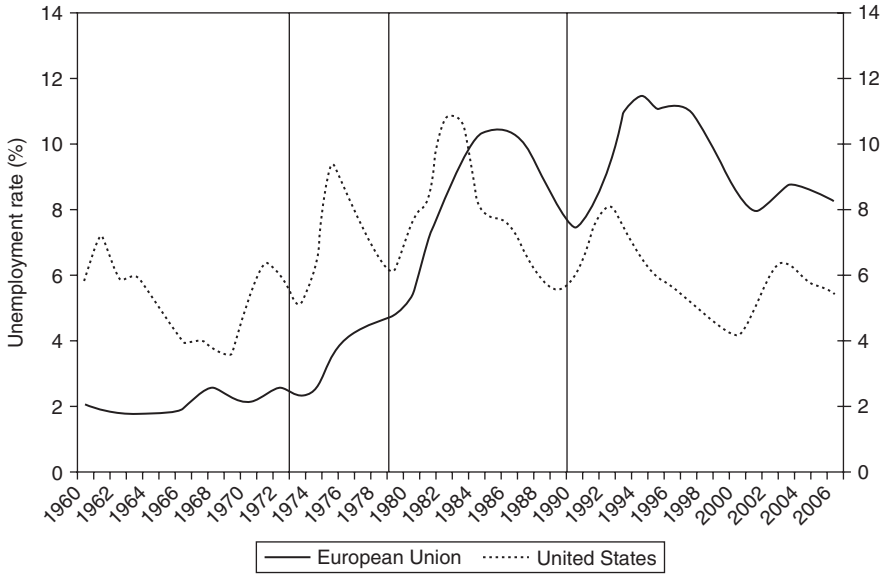


Figure 1.1 United States and European Union Unemployment Levels

global shock at the end of the 1980s (the interest rate hike) to create a sizable gap in unemployment between the two sides of the Atlantic. Up to the mid-1970s the “jobs miracle” was located in Europe, and the same “rigid institutions” currently considered responsible for European unemployment were pointed out by the U.S. literature as one of the main factors behind the European success story. For instance, in 1964 a U.S. policy-maker, Robert Myers wrote in a report that he was “looking enviously at our European friends to see how they do it” and inviting everybody to take a look at institutions on the other side of the Atlantic: “it would be short-sighted indeed to ignore Europe’s recent success in holding down unemployment.” In this initial chapter we cope with these two shortcomings by offering (1) a definition of labor market institutions acknowledging the fact that they play some useful function and (2) a simple framework explaining why these institutions, once considered a success story, are now treated as devils.

1.1 A Few Key Definitions

It is useful to start with a few key definitions that will be used henceforth.

- A *labor market* is a market where a quantity of labor services, L , corresponding to tasks specified in an *unfilled* assignment or job description (*vacant job*), is offered in exchange for a price or remuneration, called *wage*, w . Not all labor services offered by an individual are paid. For instance, the time we devote to cleaning our own apartment is not paid. It becomes market work

only if we hire a housecleaner. In order to be in the labor market, there must be an exchange of a labor service for a wage.

- The value of a job, y , is the value of the labor product obtained when the firm and the worker engage in production. One can think of it as the revenues from the job, that is, the product of the quantity of output produced by that job and the price of this output. Both the value of a job and the price of the good produced by this job may not be fixed, but may vary with the quantity of jobs and output. Thus we will typically refer to the value of the *marginal* product of labor, that is, the price of the good times the increase in output made possible by hiring an additional worker.
- The *worker's surplus* is the difference between the wage actually earned by the worker and that worker's *reservation wage*, w^r , that is, the lowest wage at which the worker is willing to accept a job offer. The reservation wage is defined as the wage that makes the worker indifferent between working and not working. Any wage earned above this level represents a net gain over the option of not working, or a *surplus* from the standpoint of the worker. Formally, the worker's surplus is given by $(w - w^r)$.
- Similarly, the surplus of the firm, also called the *firm's marginal profits from the job*, is the difference between the value of a job (the revenues from the job) and its costs, notably the wage paid to the worker engaged in that job, that is, $(y - w)$.
- The *total surplus* from a job is the sum of the worker's and the firm's surplus, that is, $(w - w^r) + (y - w) = y - w^r$. Notice that the wage, the value of a job, and the reservation wage can all be expressed in monetary terms, for instance, in euros. Hence, given y , w , and w^r , one can readily obtain the worker's surplus, the firm's surplus, and the total surplus.
- A *labor market institution* is a system of laws, norms, or conventions resulting from a *collective* choice and providing constraints or incentives that alter *individual* choices over labor and pay. Single individuals and firms consider the institutions as given when making their own individual decisions. To give an example, an individual has limited choice over the number of hours of work to be supplied when working time is determined via a collective choice mechanism. As discussed in chapter 5, regulations on working hours are an institution aimed, inter alia, at coordinating the allocation of time to work, leisure, or home activities across and within households. Because of their foundations in collective choices, institutions are the by-product of a political process. Often, institutions are established by laws, but not always. For instance, collective bargaining institutions (chapter 3) are most frequently regulated by social norms and conventions rather than by laws.

1.2 A Competitive Labor Market

These definitions suggest that it is wages that split the value of a job between the firm-employer and the worker. Wages are themselves an outcome of the labor market mechanism, allocating workers to jobs. And institutions interfere with this mechanism.

It is instructive to start by considering a *perfectly competitive labor market*, that is, a *transparent* market, where workers and firms are perfectly informed about wages and labor services offered by other firms and where there are no *frictions* or costs (e.g., no time related to job search, no transportation costs when going to job interviews) involved in the matching of workers and vacancies. Both assumptions, perfect information and a frictionless labor market, are rather extreme and seldom exist in modern labor markets. Nevertheless, the perfectly competitive labor market is a useful benchmark in our analysis.

1.2.1 Labor Supply and the Reservation Wage without Hours Restrictions

Individuals participate in the labor market and supply labor services if they can get some nonnegative surplus from working. This means that their reservation wage must be lower than or equal to the wage offered in the labor market. How is the reservation wage defined? Consider an individual whose utility function is defined over consumption, c , and leisure, l , which are both assumed to be normal goods: $U(c, l)$, whose partial derivatives are $U_c, U_l > 0$. The individual allocates the endowment of time, say l_0 , alternatively to work h , hours earning at the hourly wage, w or to leisure (clearly $h = l_0 - l$). Define nonlabor income (the income when working zero hours) as m and take the price of the consumption good as the numeraire (the price of c is one euro).

The *budget constraint* is given by

$$m + wh \leq c.$$

In the consumption/leisure space this constraint has a kink that corresponds to the level of nonlabor income, as depicted in figure 1.2. When $m = 0$, the budget constraint is a straight line crossing the horizontal axis at l_0 , where no hours of work are supplied and hence income to buy consumption goods is 0. To the left of that point (to the left of the kink), income grows at rate w because each additional hour of work yields an extra hourly wage.

The utility function can be graphically represented as a set of *indifference curves*. Each curve maps the combinations of consumption and leisure that yield the same level of utility to the worker. Because utility is increasing in both arguments, the indifference curves are negatively sloped: more consumption is needed

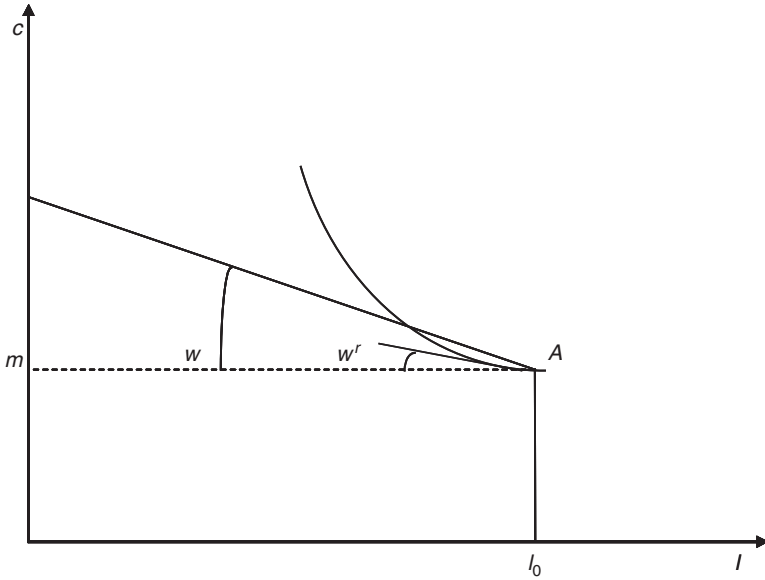


Figure 1.2 The Reservation Wage

to compensate the worker for the loss of an hour of leisure, and vice versa. The degree of convexity of these curves is decreasing with the degree of substitutability between labor and leisure. Because of our assumptions, indifference curves do not intersect, and utility is increasing farther from the origin.

The reservation wage, w^r , is given by the slope of the indifference curve crossing the kink of the budget constraint, evaluated precisely at the kink, where the individual allocates m euros to the purchase of consumption goods and works zero hours. Any wage w lower than the reservation wage will not be accepted by the individual because the marginal value of leisure (the reservation wage) exceeds its opportunity cost (the market wage). Conversely, when $w > w^r$, as in the figure, the individual who is maximizing utility will work some hours and devote the remaining time to leisure.

This definition of the reservation wage applies to conditions where the individual can choose freely how many hours to work and how many hours to devote to leisure. In real life individuals rarely have unconstrained choice of h . They have, at best, some leverage in deciding among a subset of possible hours of work, for example, between full-time and part-time jobs. This is because there is an institution (mandatory working-time legislation or collective bargaining agreements regulating working hours) that imposes, via a collective choice mechanism, constraints on individual decisions.

Box 1.1 The Reservation Wage with and without Hours Restrictions

When there are no constraints on the choice of hours, the reservation wage is given by the condition

$$\left(\frac{U_l}{U_c} \right)_A = w^r,$$

where U_l and U_c denote the marginal utility of leisure and consumption, respectively, and their ratio is the marginal rate of substitution between consumption and leisure. The rate is evaluated at zero hours of work, where the individual is buying consumption goods by drawing only on nonlabor income.

An individual free to choose how many hours to work equates the marginal rate of substitution to the market wage. Hence, when $w^r = w$, the individual is indifferent between working and not working. When $w^r < w$, the optimal choice of hours, h^* , is greater than zero. When $w^r > w$, $h^* = 0$.

Consider now a constrained choice. Suppose for simplicity that individuals actually have no choice over working hours and can only work h_{ft} hours, corresponding to a full-time job. The reservation wage will now be implicitly defined as the wage that would make the individual indifferent between not working at all and working exactly h_{ft} hours, that is,

$$U[m + w_{ft}^r(l_0 - h_{ft}), l_0 - h_{ft}] = U(m, l_0).$$

The interpretation of this condition is that when $w = w_{ft}^r$, the constrained choice is on the same indifference curve that intersects the zero hours locus. In other words, the individual is indifferent between working exactly h_{ft} hours and not working at all.

The reservation wage with hours restrictions no longer coincides with the slope of the indifference curve at the kink of the budget constraint (see box 1.1). The reservation wage with hours restrictions can be graphically represented as the slope of the segment going from the kink of the budget constraint to the locus where the indifference curve through the (m, l_0) pair crosses the vertical hours constraint, as depicted in figure 1.3. This hours-constrained choice yields a lower level of utility than the unconstrained choice, provided that the latter, at the market wage, involves some positive amount of hours of work; otherwise the hours constraint is not binding.¹

More important, the reservation wage of an individual who is constrained in terms of hours of work (w_{ft}^r) is higher than the reservation wage of an individual free to choose hours of work (w^r). Because of the concavity of the utility function,

¹ The reasons why hours are regulated although such institutions apparently reduce the well-being of an individual are discussed in chapter 5.

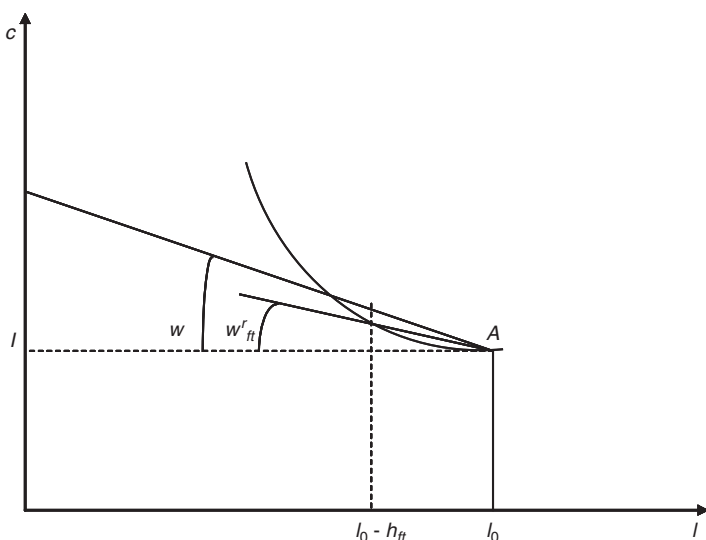


Figure 1.3 Reservation Wage with Hours Constraint

the slope of the indifference curve increases as we move to the northwest along the same indifference curve. The labor supply decision of the individual will now obey a simple rule: supply h_{ft} hours if $w \geq w_{ft}^r$ or do not offer labor services (supply zero hours) otherwise.

1.2.2 Aggregate Labor Supply

Consider now a plurality of individuals who may well have different preferences about consumption and leisure and varying endowments of nonlabor income. The reservation wage will then vary across individuals depending on their nonlabor income, as well as their preferences about leisure and work. As discussed in chapter 7, time spent outside work can also be devoted to (unpaid) activities such as household tasks generating goods and services that increase the welfare of the household. For instance, some workers may have child care responsibilities, which increase their reservation wage.

Denote by $G(w)$ the fraction of individuals of working age with a reservation wage equal to or lower than w . By multiplying this fraction by the number of persons of working age, we obtain the *aggregate labor supply* schedule. Insofar as work involves some effort, the percentage of individuals willing to work will be increasing with the wage offered to them. This means that we expect $G(w)$ to be monotonically increasing with w . By construction, $G(w)$ will also take values only in the interval bounded from below by 0 (nobody is willing to take the job at a wage lower than the lowest reservation wage) and above

by 1 (when nobody of working age has a higher reservation wage). It is certainly possible that more than one individual has the same reservation wage, in which case aggregate labor supply will involve some flat segments. It is also plausible that some individuals, for example, a rich heiress, would not work whatever the wage offered to them. Box 1.2 explains how the aggregate labor supply function can be obtained from data on self-reported reservation wages of individuals.

Box 1.2 Stated Reservation Wage and Aggregate Labor Supply

It is instructive to obtain the $G(\cdot)$ distribution, which plays a crucial role in this book, from actual labor market data. Many surveys, such as labor force surveys, in several OECD countries ask respondents about the lowest wage at which they would be willing to take a full-time job offer. This reported reservation wage is an empirical proxy for our w^r . Longitudinal data (observations of the same individuals at different times) suggest that respondents take this question quite seriously. For instance, individuals observed to be unemployed at a given date and employed at the time of the next interview generally work at a wage that is not lower than the reservation wage stated in the first place. (Needless to say, it is possible that individuals revise downward their reservation wage when they perceive that their human capital is depreciating or no longer have family responsibilities, but this does not seem to happen very frequently.) Thus individuals appear to follow consistently a reservation wage policy in their labor supply decisions (they accept only jobs offering $w \geq w^r$).

In order to obtain the $G(\cdot)$ function, we need to select individuals with similar characteristics, hence facing the same labor demand, and then count those having a reservation wage lower than any given potential wage. Because many different variables affect the productivity of workers and hence their reservation wages, it is preferable to isolate the component of each stated reservation wage that is independent of differences in productivity and has to do with differences in preferences for leisure or in the value of the time devoted to activities at home. Technically this can be done by estimating a reservation wage function, that is, a function placing premiums or discounts on observable characteristics of individuals that are valued by market wages, and then taking the differences between the reported w^r and the $\hat{w}^r$ obtained by attributing wage premiums (or discounts) according to these estimates for each individual.

Figure 1.4 displays the fraction of individuals having a residual reservation wage ($w^r - \hat{w}^r$) greater than or equal to any (conditional) wage level displayed in the vertical axis in Germany in 2001. Notice that there are various flat segments in the curve (individuals having the same residual reservation wage), and at some point the curve becomes vertical, denoting individuals who would not work whatever the wage. In order to obtain the total number of individuals wishing to work at any given wage, or the *aggregate labor supply*, we need to multiply the fraction

(continued)

Box 1.2 (continued)

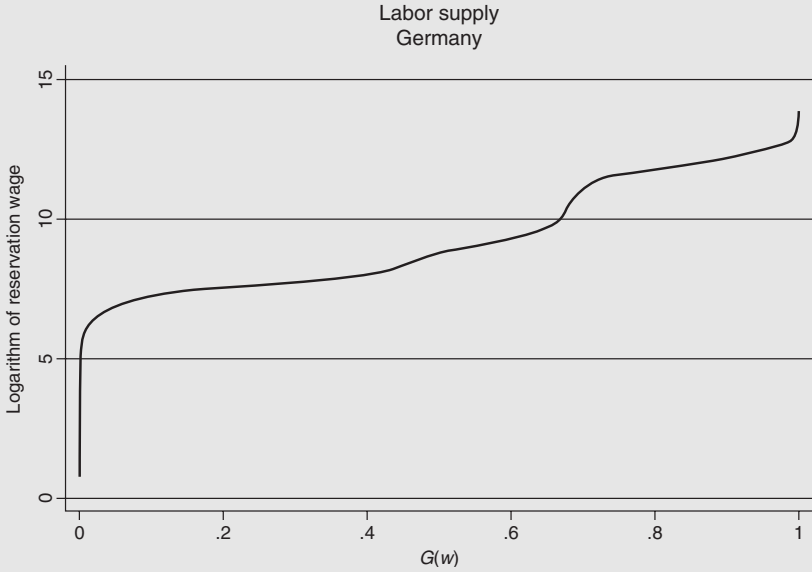


Figure 1.4 An Empirical Estimate of the Aggregate Labor Supply Function

$G(w)$ by the total population of working age, N . Insofar as N is independent of w , the aggregate supply $L^s(w)$ will have the same properties as the empirical distribution $G(w)$; notably it will be increasing with wages. A Stata data file with the primary reported reservation wage and a program (do file) generating the aggregate labor supply for the EU-15 are available on our webpage (see link at <http://press.princeton.edu/titles/8771.html>).

1.2.3 Labor Demand and Wage Determination

Production takes place by combining labor with capital. In the short run, capital is fixed so that there is no possibility to substitute labor with capital. Suppose that there is only one type of worker from the standpoint of a firm; that is, labor is homogeneous.² A profit-maximizing firm will hire workers up to the point

² Notice that we could as well assume that workers differ in terms of productivity, but that these differences are fully offset by wage differentials so that each employer is indifferent between hiring a high-productivity or a low-productivity worker.

where y , the value of the marginal job, equals the marginal cost of labor, that is, the wage. In a competitive market all firms will take this wage as given. Hence all firms will also have the same y at the equilibrium, and the *aggregate labor demand* will simply add up the number of jobs in each firm, yielding the same y . To put it another way, y provides the *marginal willingness to pay* of firms for labor services or their *inverse labor demand* schedule, $y(L)$. In order to obtain labor demand, we simply have to substitute y with w and solve for L . Formally, we set $y(L) = w$ and solve for L , obtaining $L^d(w)$.

Can we say anything about the slope of this labor demand function? By the law of diminishing marginal returns, the marginal product of labor is declining with the number of jobs for each individual firm. If not only the labor market but also the product market is competitive, then each firm will sell the product of labor at a given price, independently of the level of output. In this case the labor demand function will have the same slope as the (declining) marginal productivity of labor; that is, it will be decreasing with L , the quantity of labor being used. If instead firms have some monopoly power in product markets, the value of the marginal product of labor will include an additional term that reflects the change in price associated with the extra output produced by the additional job, multiplied by total output.³ Intuitively, when a firm faces a downward-sloping product demand curve, increasing production lowers prices of all units being sold. The less competitive the product market, the stronger the decline in prices associated with an increase in the quantity of jobs and output. By the same token, more competition in product markets involves a flatter labor demand curve.

To summarize, independent of the product market structure, labor demand, L^d , will be declining with wages, or the inverse labor demand, $y(L)$, will be declining with L . When product markets are noncompetitive, labor demand will be steeper with w .

³ Formally, for a competitive firm (superscript c), the value of the marginal product of labor (VMP)

$$VMP^c = pf_L,$$

where p is the (given) price at which output can be sold and f_L is the marginal product of labor. For a firm operating in a noncompetitive product market, we have instead

$$VMP = pf_L + p_L f_y,$$

where p_L is the marginal effect on prices of the increase in the quantity produced by the firm associated with the use of an additional unit of labor, from which it follows that $VMP = VMP^c$ when $p_L = 0$; that is, the firm is price-taker also in product markets. Because p_L is negative, labor demand of a monopolist will always be to the left of the demand curve of a competitive firm. Notice further that the difference between y^c and y is increasing in f , hence in the amount of labor being used. Thus labor demand of a monopolist will be steeper than the labor demand of a competitive firm.

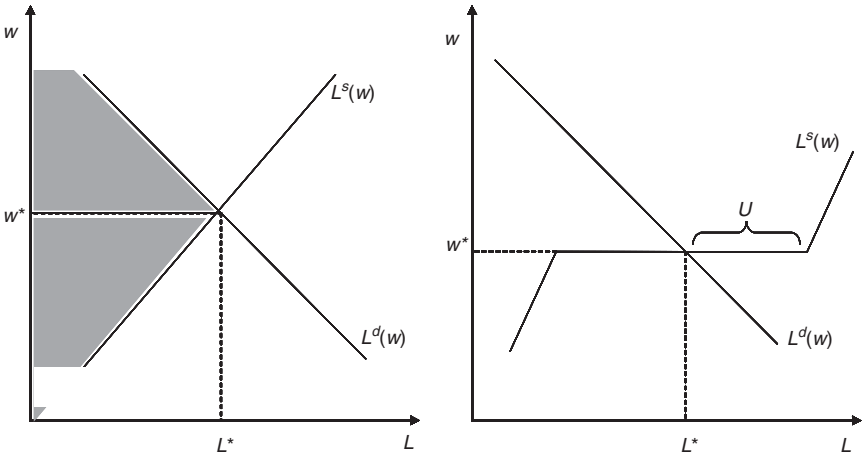


Figure 1.5 Equilibrium in a Competitive Labor Market

1.2.4 Equilibrium

Figure 1.5 depicts a downward-sloping labor demand together with an upward-sloping aggregate labor supply. In a competitive labor market the equilibrium wage level, w^* , will lie at the intersection of the two curves. It is important to notice that there is only one wage level being determined at the equilibrium. This means that workers with a reservation wage strictly lower than w^* will realize a positive surplus from participating in the labor market. The sum of all these individual surpluses is given by the shaded area below the wage and above the labor supply. Firms will also realize some surplus or profits. This is depicted as the shaded area above the equilibrium wage and below the labor demand schedule.

Workers with a reservation wage larger than w^* will instead decide not to work. In other words, $L^* = G(w^*)$ will be the *employment rate* (the fraction of the working-age population holding a job), while $1 - G(w^*)$ will be the equilibrium *nonemployment rate*. Notice that the equilibrium wage level may well be in a flat segment of the labor supply curve. In this case there will be individuals with $w^r = w^*$ who are not working even if they are willing to work at the equilibrium wage. These individuals are, strictly speaking, *unemployed*, as denoted by the segment U in the right-hand panel of figure 1.5, although they do not suffer any welfare loss from not working (the fact that $w^r = w^*$ means that they are indifferent between working and not working). All other nonemployed individuals are *inactive* according to internationally accepted definitions of labor market status (see box 1.3 for a discussion of these definitions, notably the borders between unemployment and inactivity).

Box 1.3 Definitions of Labor Market Status and the Porous Participation Borders

How precisely is the labor market status of individuals defined by available statistics? According to internationally accepted (OECD–International Labor Organization (ILO) definitions, the entire population of working age (15–64) can be classified in three main labor market conditions: individuals can be either employed, unemployed, or inactive. An *employed individual* is someone in the armed forces or who has worked for pay (in cash or in kind) at least one hour during the reference period (a week or a day) or has a formal attachment to a job but is temporarily not at work (e.g., because of an illness, a holiday, or maternity leave). A person of working age is classified as an *unemployed individual* if the following five conditions are fulfilled:

1. The person is currently not working.
2. The person has looked for work in the four weeks before the survey.
3. The person has looked for work actively (e.g., sending applications to employers or contacting a private placement agency or a public employment office).
4. The person is willing to work.
5. The person is immediately available for work, meaning that the person can start a job within two weeks following the interview.

Finally, *inactive individuals* are persons who are neither employed nor unemployed according to these definitions. This residual group includes a highly heterogeneous population. Consider, in particular, the borders between unemployment and inactivity, which are very important in determining the size of labor supply. There are individuals not working who do not satisfy any of the five conditions for unemployment and individuals who fail on just one account, such as one who has been actively looking for work and is willing to work, but is not immediately available because of a temporary disability. Similarly, a reduction in search intensity, such as the fact of not having sought a job actively in the reference period, implies that one person has moved from unemployment to inactivity even if that person is still looking for jobs, willing to work, and immediately available. All this suggests that unemployment statistics may exclude an important component of labor supply.

Thus it is always a good idea to go beyond these definitions and adopt broader measures of potential labor supply. The OECD pioneered this work by developing *supplementary measures of labor slack*, which added to the official unemployment rate statistics *discouraged workers* (persons failing the job search requirement because they feel that no suitable job is available for them). As suggested in OECD (1995), in some countries discouraged workers account for more than 2 percent of the labor force; hence their inclusion in unemployment statistics could significantly increase these official (and highly politically sensitive) measures of labor slack.

(continued)

Box 1.3 (continued)

Measures based on OECD-ILO definitions (% of working age population), 1994–2000						
Country	Empl.	Unempl.	Out of the Labor Force			
			Total	Potentials	Discouraged	Unattached
Austria	67.5	2.9	29.6	1.0	0.3	28.3
Belgium	59.0	4.1	37.0	1.7	0.7	34.6
Denmark	74.1	4.9	21.0	3.4	0.4	17.2
Finland	63.4	8.4	28.2	3.5	0.4	24.3
France	60.6	7.0	32.4	1.7	0.1	30.6
Germany	64.5	5.7	29.8	1.3	0.2	28.3
Greece	54.5	6.7	38.9	0.8	0.5	37.6
Ireland	58.1	6.5	35.4	1.2	1.3	32.9
Italy	51.8	8.6	39.7	2.8	0.5	36.4
Luxembourg	61.2	2.7	36.1	0.8	0.0	35.3
Netherlands	67.8	3.9	28.3	1.1	0.1	27.1
Portugal	67.7	3.6	28.6	1.5	0.2	26.9
Spain	48.3	12.4	39.4	1.7	0.3	37.4
United Kingdom	68.8	7.4	23.8	1.2	0.3	22.3

Jones and Riddell (1999) analyzed the behavior over the cycle of the OECD-ILO unemployed and of these broader measures at labor slack in Canada and concluded that there are significant comovements in these series. More recently, statistics of the *potential labor force* have been developed by the European Commission that include persons who are willing to work but have not been actively searching in the previous four weeks. In particular, a report by the European Commission (2005) suggests that these “potential members of the labor force” account for almost 15 percent of the OECD-ILO inactive in the EU-25.

A good test of whether these potential members of the labor supply are closer to inactivity or unemployment is provided by analyzing transitions from nonemployment into employment. Brandolini, Cipollone, and Viviano (2006) suggest that the potential labor force group is a somewhat intermediate state between unemployment and inactivity: it has lower chances of getting a job than the OECD-ILO unemployed, but higher chances than the other inactives. However, within these potential members of the labor force there is a smaller group of individuals who are indistinguishable from the unemployed in terms of their labor market transitions. As suggested by the table in this box, drawn from their study, “potential members of the labor force” and “discouraged workers” account in some countries for up to 4 percent of the population of working age. Corrections of unemployment rate statistics based on the OECD-ILO definitions by using the “virtually unemployed” inactive, as defined earlier, could increase the unemployment rate in Italy by almost two percentage points.

Sources: Brandolini, Cipollone, and Viviano 2006; European Commission 2005; Jones and Riddell 1999; OECD (1995).

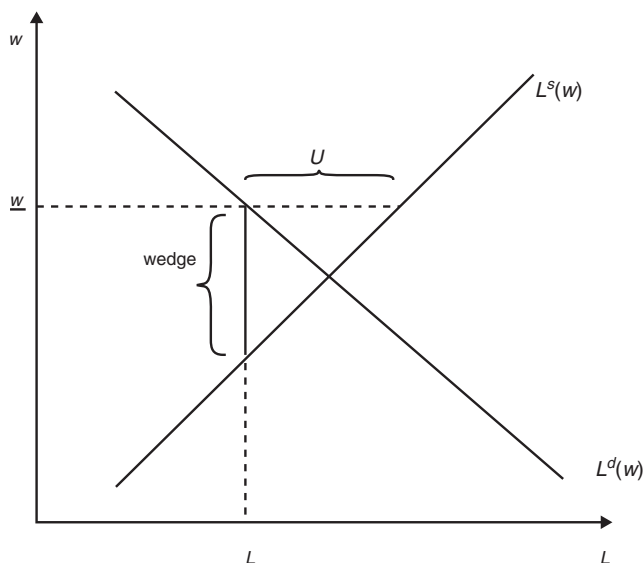


Figure 1.6 Price-Based Institutions and the Wedge

1.3 Labor Market Institutions

We are now ready to describe how *labor market institutions* operate. According to our definition, they are outcomes of collective choice mechanisms that interfere with the exchange of labor services for pay. They do so by introducing a *wedge* between the reservation wage of the workers and the value of a job, that is, between the labor supply and labor demand schedules.

1.3.1 Acting on Prices

Let us give a few examples of how labor market institutions operate. Formal descriptions are provided in technical annex 1.4. An institution like the *minimum wage* (chapter 2) sets a lower bound $\underline{w}$ to the wage paid to individual workers. By doing so, it changes the slope of the labor supply schedule, preventing employer-firms from hiring workers at a lower wage than the minimum wage even when the reservation wage of those supplying labor services is lower than $\underline{w}$. The actual labor supply faced by employers is now represented by the dotted line in figure 1.6. The latter coincides with the reservation wage schedule only to the right of $L^s(\underline{w})$. Notice further that the segment $L^s(\underline{w}) - L^d(\underline{w})$ denotes *unemployed* individuals, that is, persons who are *not working*, but would be *willing to work* at the equilibrium wage. Insofar as their reservation wage is lower than $\underline{w}$, these individuals will not be indifferent between working and not working. In other words, unlike

in a competitive and institution-free labor market, we now have strictly a welfare *loss* associated with unemployment.

There are various ways to implement a minimum wage. In some countries there is a statutory minimum wage set by the government. In other countries a *trade union* (chapter 3) imposes floors for wages via collective wage agreements in specific industries. Collective bargaining is itself an institution that interferes with wage setting not only by setting minima for pay, but also by affecting wages above these minima, for example, by imposing egalitarian wage scales. When unions are present in the workplace, employers face a labor supply schedule that departs from the reservation wage of each individual worker. Unions thus impose on employers the payment of a *markup* over the reservation wage of individuals.

Taxes on labor (chapter 4) are another institution that introduces a wedge between reservation wages and the value of labor productivity. They reduce labor demand, but also labor supply because some individuals drop out of the labor force who would be willing to work in the absence of taxes. This means lower employment and participation, but no unemployment unless the net wage happens to be in a flat segment of the labor supply schedule.

The proceeds of labor taxes are generally used to finance *retirement plans* (chapter 6), *family allowances* (chapter 7), and *unemployment benefits* (chapter 11). All *nonemployment benefits* (subsidies provided conditional on not working) shift labor supply upward, reducing the employment rate and the size of the labor market. Part of this reduction in employment is accommodated via an increase in unemployment, and the remaining part via an increase in inactivity. The magnitude of the effects on inactivity and unemployment depends on the institutional details, notably on whether payments are contingent on nonemployment or require some job search effort (e.g., unemployment benefits may be accompanied by the activation measures outlined in chapter 12, which implement work tests, eliciting job search effort, for those receiving the benefits).

1.3.2 Acting on Quantities

Minimum wages, trade unions, taxes, and unemployment benefits operate mainly on the price of labor side. They *directly* introduce a wedge between y and w' by forcing employers to pay more than the reservation wage of the marginal job and workers to receive less than the labor cost paid by employers. Other institutions act on the quantity of labor being supplied or demanded and hence introduce a wedge only *indirectly* because the actual or *effective labor supply* faced by employers departs from the cumulative distribution of individuals' reservation wages.

For instance, *regulations on working hours* (chapter 5), *immigration policies* (chapter 9), or an increase in the *compulsory schooling age* (education policies are discussed in chapter 8) cut away a segment of the population of working age. It is plausible that most of the individuals who can no longer supply labor under these

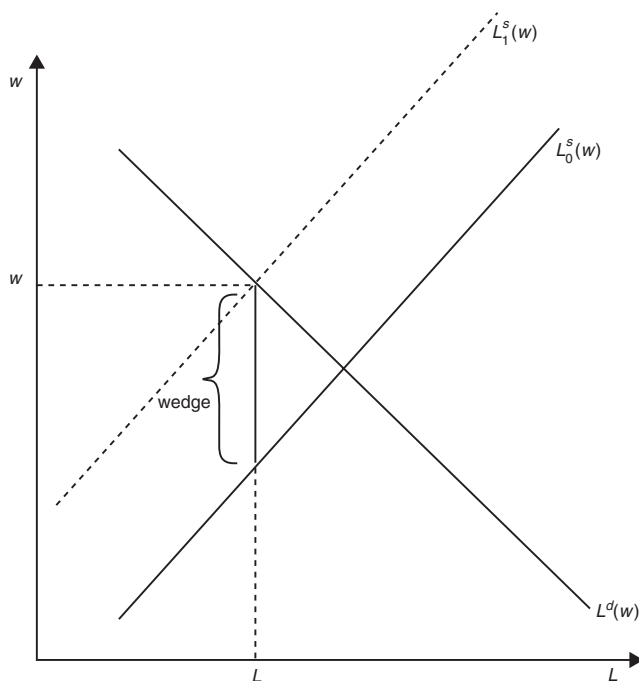


Figure 1.7 Quantity-Based Institutions and the Wedge

restrictions (e.g., women after maternity, first-time jobseekers, and migrants) have a relatively low reservation wage; that is, for them, $w^r < w^*$. Thus the quantitative restrictions cut away a segment of labor supply to the left of w^* , involving a shift to the left of the entire schedule, as depicted in figure 1.7. The new equilibrium will feature higher wages and less employment, just as in the case of institutions that act on prices. Once more, labor market institutions operate by introducing de facto (in this case indirectly) a wedge between the value of the marginal product of labor and the reservation wage. By reducing the segment of the population for which $w > w^r$, they reduce not only the working-age population, but also the size of the labor force and the employment rate.

Another common quantity restriction in industrialized countries is *employment protection legislation* (EPL) (chapter 10). This legislation makes it more costly for employers to adjust the number of workers in a firm in response to shocks. Unlike payroll taxes, EPL involves taxes and transfers to workers that are paid only in case of dismissals. Employers must pay social security contributions to employ labor, and they reduce employment in the face of higher payroll taxes if labor demand is downward sloping. But they can avoid paying firing costs by choosing a stable employment path around a level that may be slightly lower or even higher on average than what would obtain, for the same wage and contributions level, in the absence of EPL. This does not imply that firms should be

happy to do so: by definition, whenever firms fail to equate w to y , they earn lower profits. In this sense it is quite reasonable to think of employment protection legislation as imposing a tax on employers. Still, EPL does not reduce profits through lower average employment levels, but rather through poor synchronization of productivity and wages around roughly unchanged average levels.

1.3.3 Institutional Interactions

As just argued, EPL that imposes dismissal costs acts mainly on labor market *flows*. It does so by reducing the incentives for firms to shed labor. It is perhaps a little less intuitive that EPL also reduces incentives to hire: if employers anticipate that layoffs will be difficult or costly, in fact, they should try to reduce the amount of labor shedding called for by future labor demand downturns or wage upturns. This means hiring fewer people from the start. Because both firings and hirings decline, the net effect on employment and unemployment levels is ambiguous.

Yet EPL may indirectly affect employment by giving more power to trade unions in wage bargaining, and in this case the impact is likely to be unambiguous. Stronger bargaining power of workers shifts the labor supply faced by employers upward, increasing the equilibrium wage and reducing aggregate employment. In other words, EPL negatively affects employment by interacting with other institutions, such as collective bargaining institutions.

These institutional interactions can be complex, and there can be many of them, given that there are several possible combinations of institutions in place. In chapter 13 we discuss the interactions that appear to us most relevant. Unavoidably the list is not exhaustive. The important thing to remember at this stage is that one should never confine the analysis to the simple direct effect of one institution on the labor market. We live in labor markets in which institutions never operate in isolation.

It is customary to describe the institutional landscape of Organization for Economic Cooperations and Development (OECD) countries in terms of a *cluster of institutions*. For instance, the so-called Nordic model (Denmark, Sweden, Finland, and the Netherlands) features generous nonemployment benefits combined with rather strict activation policies and the involvement of unions in the administration of unemployment benefits. Another example is the Southern model (including Italy, Greece, Spain, and Portugal), which features traditionally relatively strict employment protection legislation, early retirement provisions, and a rather strong influence of trade unions.

These different clusters of institutions involve very different labor market outcomes. As shown in figure 1.8, there is wide variation across OECD countries in employment rates (the ratio of employment to the population of working age) and unemployment rates (the ratio of unemployed workers to the labor force). Another fact illustrated by figure 1.8 is that the same employment rate can be achieved at

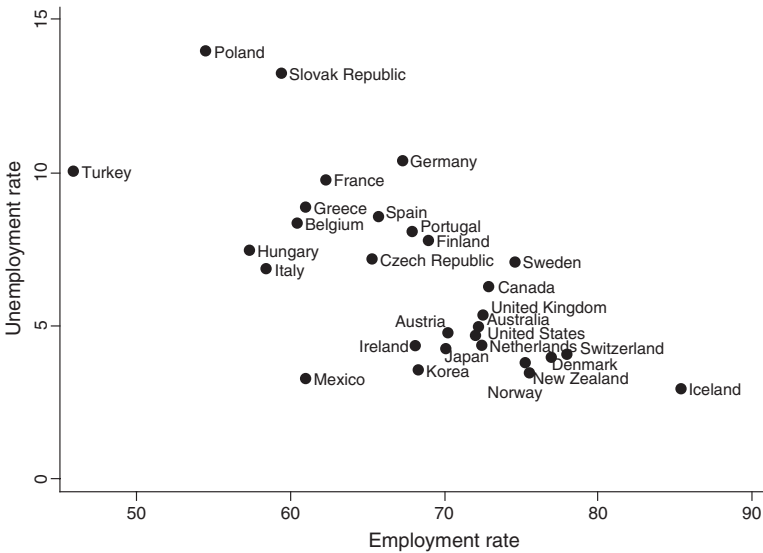


Figure 1.8 Employment and Unemployment Rates in OECD Countries (2006)

less than 5 percent or at double-digit unemployment rates. This suggests that it is important not to neglect labor force participation effects. Mexico, for instance, has an employment rate 20 percent lower than Iceland but has the same unemployment rate, as many more individuals of working age are inactive.

1.3.4 Why Do Labor Market Institutions Exist?

Because all labor market institutions introduce a wedge between labor demand and supply, they reduce the size of labor markets. If the labor market is competitive, the total surplus to be shared between firms and workers will be reduced after the introduction of any labor market institution. The obvious question is why these institutions are so important in modern labor markets. They are certainly not imposed by heaven. They are introduced by democratically elected governments. If voters did not like these institutions, they would sooner or later be removed. If these institutions reduce the size of the economic pie, then it should be possible to make everybody happier (or at least as happy) without them.

We offer three arguments for the existence of labor market institutions:

1. *Efficiency.* A first-best competitive labor market outcome is unattainable; there are second-best arguments justifying the presence of these institutions.
2. *Equity.* In the absence of nondistortionary taxes and transfers, these institutions are best suited to achieve some redistribution that is supported by voters.

3. *Policy failures.* There are failures in the political process that make it possible for minority interest groups to succeed in imposing their preferred institutions on majorities who would be better off without them.

Often these three reasons coexist, but we discuss them separately for the sake of simplicity. We confine ourselves here to a few illustrations of how these mechanisms operate. Later chapters contain a thorough discussion of the rationale for each institution.

Efficiency

Labor market institutions exist because there are *market imperfections* that prevent the institution-free equilibrium from attaining the competitive equilibrium outcome. In practice, a competitive market for labor does not exist. Labor markets are far from competitive because there are important informational asymmetries between employers and employees, as well as externalities, that is, goods produced and consumed that are not subject to market interactions. In both cases—*asymmetric information* and *externalities*—labor markets violate the transparency and complete market properties of a competitive market. Well-designed labor market institutions, in this context, may remedy these failures of markets and increase the size of the pie compared with the *laissez-faire* outcome.

Equity

Institutions change the allocation of the surplus between employers and employees. Even when they reduce the size of the economic pie, they can make one side of the market (those supplying labor services or those buying them) strictly better off than without the institutions. In principle, redistribution could also be achieved by taking the *laissez-faire* outcome and then taxing employers or employees and transferring the proceeds to the other side of the market. In practice, however, redistribution via *lump-sum taxes and transfers* is not possible because redistributive policies can only rely on information, on signals, that can be altered at will by individuals. This means that any type of redistribution is unavoidably distortionary and labor market institutions, such as distortionary labor taxes and transfers, can be the most efficient way to redistribute.

Policy Failures

Because of these redistributive properties of institutions, there are also instances in which some powerful minorities succeed in imposing a set of institutions on a majority of citizens. This happens particularly when the benefits of an institution are concentrated in a small segment of the population while the costs are spread over a very large crowd of individuals. Under these conditions, groups organized as a lobby may succeed in influencing political decisions disproportionately.

A Few Examples

In practice, labor market institutions perform several functions at once: they remedy market failures but, at the same time, affect the income distribution or meet the requests of specific interest groups. For example, in the absence of perfect capital markets, the welfare of risk-averse individuals can be increased by offering insurance against the risk of income fluctuations. Job loss is one of the occurrences against which workers could be protected. However, no private insurer will ever want to provide insurance against unemployment because *moral hazard* and *adverse selection* stand in the way of these potential contractual arrangements. Workers would not try as hard to avoid unemployment and find new jobs if they were covered against the negative consequences of the event by purchasing insurance at a given market price (moral hazard), and workers who know that their unemployment risk is particularly high would make the scheme unprofitable for insurance providers and/or unattractive to workers with average risk (adverse selection). This explains why collective action (institutions) tries to remedy the inequitable or unfair labor market treatment of workers who, lacking insurance, become or remain unemployed despite their best efforts. Unemployment benefits and employment protection legislation are remedies for this failure of markets. By supplying insurance, however, they involve some trade-offs. For instance, provision of insurance in the presence of asymmetric information unavoidably decreases productive efficiency. Workers have no less incentive to decrease their jobseeking effort when they are covered by social rather than private insurance, and protection from supposedly unfair developments unavoidably decreases the labor market's speed of adjustment.

While remedying a market failure, employment protection legislation and unemployment benefits transfer resources from employers to employees, creating a vertical redistribution of income. Most of the institutions analyzed in this book address distributional tensions by attributing a larger share of the economic pie to workers or to nonworking individuals and extracting surplus from employers. Minimum wages, restrictions on hours of work, collective bargaining institutions, and unions respond to distributional concerns by assigning a larger share of the pie to workers even at the cost of generating overall a smaller pie. At the same time, these institutions remedy market imperfections, such as the presence of monopsonistic power of firms and externalities in the wage-setting process and in bargaining over hours. Migration restrictions also have a well-defined distributional objective: they insulate native workers from competition of foreign workers. Their presence can also be explained in terms of market failures associated with interactions with other institutions. In the presence of minimum wages, migrants may crowd out native workers, or migrants who do not find a job may exert a negative fiscal externality on the native population by drawing nonemployment benefits without having contributed to their financing. Taxes on labor are often

progressive. This suggests that they pursue vertical redistribution. At the same time, however, they can be rationalized by interactions with other institutions: someone has to pay for nonemployment benefits, active labor market policies, family policies, and formal education.

In technical annex 1.4 we provide a simple formalization of the redistributive role of labor market institutions. We model a competitive market with a government caring about income distribution or agents bargaining over wages, and we obtain the optimal size of an institution. Institutions do not always have the optimal size because specific interests prevail. Strict employment protection, for instance, involves large implicit transfers from the unemployed to employees or to some categories of employees who are *de facto* insulated by employment protection from competition by *outsiders*. More broadly, the combination of price and quantity institutions that is present in many labor markets is successful in protecting *insiders* from negative labor market developments: not only are wages compressed and stable, but also tenure lengths of regular workers are clearly much longer in more rigid labor markets. Unsurprisingly, it is the insiders who oppose reforms of these institutions, even when they are a minority and when the optimal size of the wedge (operating the desired amount of vertical income distribution) would be lower. Often labor market institutions tend to privilege minority subsets of the market's labor force. Such policy failures can emerge over time as economies are hit by shocks (Blanchard and Wolfers 2000) or the economic environment is altered (Ljungqvist and Sargent 2003). The model in technical annex 1.4 suggests that the redistributive properties of institutions should be adjusted to the economic environment in which they operate. If product markets become more competitive, then redistribution involves higher costs in terms of forgone efficiency (Bertola and Boeri 2002). Under these conditions, it is better to pursue the same distributional objectives by imposing a smaller wedge between labor demand and labor supply. But policy failures may make this adjustment more difficult or altogether prevent it.

1.3.5 Product Market Competition and Institutional Reforms

As stressed earlier, it is always important to recognize that institutions fulfill a useful purpose from the point of view of at least some economic agents. Otherwise it would hardly be possible to see why they were introduced in the first place. Institutions are also subject to frequent adjustments. Table 1.1 provides information on the number of reforms carried out in the European Union (EU) in the field of labor market and social policies since 1986. It draws on the Inventory of Labor Market Reforms assembled by the Fondazione Rodolfo Debenedetti, which takes stock of reforms carried out in Europe in the field of employment protection legislation (EPL), nonemployment benefits (NEBs, encompassing not only unemployment benefits, but also the various cash transfers provided to individuals not working

Table 1.1 Number of Reforms of Labor Market Institutions in Europe (1986–2005)

		Decreasing the wedge				Increasing the wedge				Total per row	Of which decreasing (%)
		1986–90	1991–95	1996–2000	2001–5	1986–90	1991–95	1996–2000	2001–5		
EPL	<i>marginal</i>	5	7	14	12	8	7	20	18	91	42
	<i>radical</i>	–	4	4	2	–	2	1	1	14	71
NEBs	<i>marginal</i>	9	24	75	94	10	8	21	19	260	78
	<i>radical</i>	1	2	3	12	–	–	–	–	18	100
RET	<i>marginal</i>	12	22	24	25	11	9	16	16	135	61
	<i>radical</i>	1	4	5	3	1	–	–	–	14	93
WT	<i>marginal</i>	–	1	17	20	–	–	4	3	45	84
	<i>radical</i>	–	–	–	–	–	–	1	1	2	0
MIG	<i>marginal</i>	–	2	8	7	–	4	5	11	37	46
	<i>radical</i>	1	–	4	–	–	2	1	2	10	50
<i>Total per column</i>		29	66	154	175	30	32	69	71	626	68

Source: Fondazione Rodolfo De Benedetti Inventory of Labor Market Reforms (www.frdp.org).

Notes: Marginal reforms involve minor adjustments of the norms.

Radical reforms involve comprehensive adjustments of the norms.

EPL: Employment protection legislation (chapter 10).

NEBs: Nonemployment benefits (unemployment benefits, chapter 11, and active labor market policies, chapter 12).

RET: Retirement schemes (chapter 6).

WT: Working-time regulations (chapter 5).

MIG: Migration policies (chapter 9).

while they are of working age), and provisions for retirement (RET, relevant in determining participation among older workers), as well as working-time regulations (WT) and migration policies (MIG). Reforms are classified on the basis of whether they reduce (e.g., by making EPL more flexible or reducing the generosity of unemployment benefits and pensions or increasing working-time flexibility and reducing migration restrictions) or increase the wedge between labor supply and demand, as well as whether they are structural (involving comprehensive reforms of the regulations bound to involve the entire population of working age) or marginal (involving rather minor adjustments of the norms or creating dual regimes, with a minority segment of beneficiaries involved in the new regime).⁴

As documented by table 1.1, many reforms of labor market institutions are taking place. In the observation period 626 reforms were counted, that is, more than 1.2 per year and country. More than 9 reforms out of 10 are marginal. They do not necessarily go in the same direction. Most reforms decrease the wedge, but roughly 30 percent of the reforms go in the opposite direction. Often reforms undo previous reforms, by newly increasing the wedge. Due to the competitive pressures arising from globalisation, we would expect reforms to reduce the wedge. However, often reforms undo previous reforms by newly increasing the wedge. These inconsistencies and the marginal nature of most reforms point to strong political opposition to reforms: the reforms increasing the wedge are typically more popular and hence find fewer political obstacles in their way.

Table 1.1 also documents an acceleration of reforms, notably of those decreasing the wedge, in recent years. However, such reforms act mainly via marginal adjustments, as defined earlier. The ratio of marginal to structural reforms indeed has increased since 1990.

A possible interpretation of the acceleration of reforms experienced by European countries in recent years is that stronger product market competition associated with globalization increases the efficiency costs of these institutions, inducing stronger pressure for change. In particular, a more elastic labor demand brought about by increased product market competition increases the employment cost of these institutions (see technical annex 1.4). At the same time, the fact that greater competition in product markets reduces the employment levels compatible with these institutions suggests that there will be strong political resistance to downscaling the institutions that protect against labor market risk. This may help to explain why many reforms also go opposite to the direction implied by increased product market competition. Moreover, several empirical studies (e.g., Rodrik 1998; Wacziarg and Horn Welch 2003) found a positive correlation between exposure to product market competition—measured in terms of trade

⁴ Details on the inventory of social policy reforms and on each single regulatory change are offered in the webpage of the Fondazione Rodolfo De Benedetti (www.frdb.org).

openness—and the presence of redistributive institutions, pointing to stronger demand for protection in competitive environments.

Whatever the reasons for the reforms of labor market institutions, many of them occur every year. This offers a great opportunity to understand their effects on the labor market. In this book we will often refer to studies that use reforms as *policy experiments*, allowing the researcher to better isolate the effects on the labor market of any specific institution and identify the underlying causal relationships. Often not only do institutions affect labor market outcomes, but also the underlying conditions of the labor market affect the institutions. The labor market itself gives rise to political pressures to introduce, preserve, or reform these institutions.

1.4 Technical Annex: A Simple Static Framework

A simple static model originally developed by Bertola and Boeri (2002) can be valuable in characterizing equilibriums in competitive labor markets, as well as the role of labor market institutions.

1.4.1 A Competitive Labor Market

In the model below a crucial role is played by labor demand and supply elasticities, defined as the percentage change in labor demand and supply, respectively, associated with a one percentage change in the wage. On the demand side of the market, profits are maximized when the opportunity cost of a marginal job, w , is equal to its value, y . In the short run (when capital is fixed) there is no loss in generality in assuming that the marginal value of a job is a decreasing (at a constant elasticity) function of the employment rate L , that is, $y = AL^{-\eta}$, where A indexes labor productivity, and the index of the (inverse) labor demand elasticity, η , takes values between 0 (flat labor demand at A) and 1 (vertical labor demand at 1). We can then write the labor demand schedule as follows:

$$L = \left(\frac{A}{w} \right)^{\frac{1}{\eta}}. \quad (1.1)$$

The supply side of the labor market is given by the cumulative distribution function of the reservation wages, which is, by construction, increasing with w . We assume also that this schedule has a constant-elasticity functional form so that

$$G(w) = w^{\frac{1}{\epsilon}}. \quad (1.2)$$

The elasticity parameter may range between 0 (in which case the labor supply is flat and normalized to unity) and plus infinity: larger values of ϵ denote increasingly

inelastic labor supply schedules, and as ϵ tends to infinity, labor supply becomes perfectly vertical.

We consider first the equilibrium in a competitive and wedge-free labor market where $y = w^r = w^*$. By equating the two schedules, solving for L , and substituting the result in the labor supply function, we obtain

$$L^* = (A)^{\frac{1}{\epsilon+1}} , \quad w^* = A^{\frac{\epsilon}{\epsilon+1}} \quad (1.3)$$

It is easy to show that this equilibrium maximizes the total surplus from labor exchange. If we neglect irrelevant constants of integration (indexed by ξ), profits of employers are given by

$$\int_{\xi}^L Ax^{-\eta} dx - wL = \frac{A}{1-\eta} L^{1-\eta} - wL.$$

Similarly, the total surplus of workers is given by

$$wL - \int_L^{\xi} x^{\epsilon} dx = wL - \frac{L^{\epsilon+1}}{\epsilon+1}.$$

Maximizing the joint surplus (the *sum* of firm's profits and of the workers' surplus from employment),

$$\begin{aligned} \max_L \left(\left[\frac{AL^{1-\eta}}{1-\eta} - wL \right] + \left[wL - \frac{1}{\epsilon+1} L^{\epsilon+1} \right] \right) \\ = \max_L \left(\frac{AL^{1-\eta}}{1-\eta} - \frac{1}{\epsilon+1} L^{\epsilon+1} \right), \end{aligned} \quad (1.4)$$

yields the wedge-free, perfect labor market wage and employment levels (1.3). Hence the competitive outcome has the desirable property of maximizing the total surplus of production over the opportunity cost of employment, or the size of the economic pie generated by the labor market. Since maximization entails equality at the margin of the value of a job for the employer and workers' reservation wages, the competitive outcome also features no welfare loss from unemployment. Yet as long as w^* lies on a flat segment of the function $G(w)$, at the equilibrium there may be individuals unemployed, meaning in this particular case that they are indifferent between working and not working.

1.4.2 Labor Market Institutions

As discussed in this chapter, the presence of labor market institutions can be rationalized in terms of market failures as well as distributional tensions, either related to general interest redistribution in favor of workers or special interests of specific categories of workers-citizens. Market failures may arise from imperfect

or asymmetric information or because of an excessive concentration of power in the hands of employers (monopsony power), forcing both employment and wages to be lower than the optimum. Distributional concerns may arise with and without market failures. In the absence of lump-sum redistribution, even equilibriums that maximize the joint surplus (the equilibrium in a competitive economy) do not necessarily address distributional tensions within the economy.

1.4.3 The Wedge

All labor market institutions operate by introducing a wedge between labor supply and demand. Their rationale can be illustrated by comparing the institution-free, laissez-faire equilibrium with the solution of a normative social planning problem involving a choice of the size of this wedge. If the wedge is zero, the social optimum coincides with the laissez-faire equilibrium, and there is no role for labor market institutions. The size of the wedge measures the deviation of the social optimum (or the equilibrium imposed by bargaining over the distribution of the surplus) from the laissez-faire equilibrium.

In particular, consider an institution that introduces a wedge between labor supply and demand through a proportional tax on labor income, t , whose proceeds are given to the workers. Suppose that the normative criterion is the maximization of a Bernoulli-Nash social welfare function of the type

$$W = \max \left(\left[\frac{AL^{1-\eta}}{1-\eta} - w(1+t)L \right]^{(1-\beta)} \left[w(1+t)L - \frac{1}{\epsilon+1} L^{\epsilon+1} \right]^\beta \right), \quad (1.5)$$

where the parameter β measures the distribution weight of labor, that is, the importance given by society to the (functional) share of the pie going to workers. Conversely, $(1 - \beta)$ is the distribution weight of employers, and t , labor taxes, is the control variable.

By deriving the first-order condition with respect to t , imposing that we are on labor demand (1.1) and solving for the wedge, we obtain

$$(1 + t^*) = \frac{\beta}{1 - \beta} \frac{\epsilon + 1}{\epsilon} \frac{\eta}{1 - \eta}, \quad (1.6)$$

where t^* is the optimal tax on labor income. When $t^* > 0$, the social optimum involves a positive tax on labor, introducing a wedge between labor supply and demand.

It is instructive to consider the wedge-free case. Equation 1.6 suggests that $t^* = 0$

$$\mu = \frac{1 - \beta}{\beta} = \frac{\eta}{1 - \eta} \frac{1 + \epsilon}{\epsilon} \quad (1.7)$$

where μ is any wedge introduced by labor market institutions.

In other words, the ratio of the distribution weight of employers to that of workers should equal the product of the labor demand and supply elasticities. The larger the η , the lower the elasticity of labor demand and the higher the distributional weight of employers justifying a laissez-faire equilibrium. Analogously, the larger the ϵ , the lower the elasticity of labor supply and the lower the distribution weight of employers justifying a laissez-faire equilibrium. The economic intuition behind these results is that, in line with optimal taxation theory, it is better to tax more the less elastic side of the market, maximizing tax revenues. Only a strong distributional concern of this less elastic side of the market can move the equilibrium away from this optimal taxation rule.

Importantly, there is no reason to expect a priori that condition (1.7) is satisfied because β bears no systematic relationship to labor demand and supply elasticities. To put it another way, it can only be by chance that (1.7) is satisfied. In the general case, when distributional concerns are relevant, it is optimal to have some wedge between labor supply and demand, even at the cost of deviating from the equilibrium that maximizes the joint surplus. Redistribution is one of the key functions of the labor market institutions discussed in this book. The other cases for labor market institutions arise when the laissez-faire equilibrium does not maximize (1.4), and hence labor market institutions do not necessarily involve an efficiency-equity trade-off.

1.4.4 Product Market Competition and the Employment Bias of Institutions

Notice that the distribution weight compatible with the competitive, laissez-faire equilibrium is decreasing with the elasticity of demand and supply. By the same token, the *disemployment bias* of labor market institutions (the reduction in employment induced by the wedge with respect to the competitive outcome) is larger in the presence of a larger elasticity of demand. In particular, by denoting by the superscript I the presence of some institution, the disemployment bias is given by

$$\mu = \left(\frac{1 - \eta}{1 + \epsilon} + \beta \frac{\eta + \epsilon}{\epsilon + 1} \right) \frac{\eta}{1 - \eta}, \quad (1.8)$$

where μ is the markup imposed by institutions over the competitive wage.

This suggests that the equilibrium with institutions involves lower employment than at the laissez-faire competitive equilibrium when the markup is greater than 1.⁵

⁵ When the markup is strictly lower than 1, it is labor supply to be the short side of the market. Also in this case there is less employment than at the competitive equilibrium.

Suppose now that labor demand becomes more elastic, for example, as a result of a globalization shock involving greater competition in product markets. Insofar as labor market institutions do not automatically adjust to the changes in the economic environment, the employment levels before and after globalization (denoted by the subscript G) are given by

$$\beta L_G^I = A\mu_0^{-\frac{1}{\varepsilon+\eta_G}} < L^I = A\mu_0^{-\frac{1}{\varepsilon+\eta}}.$$

Thus, if the wedge remains at its optimal level (μ_0) before the globalization shock and does not adjust to the changes in the labor demand elasticity parameter, an increase in product market competition leads to lower employment, and by (1.8) there is a larger employment bias of labor market institutions with respect to the laissez-faire outcome in a competitive labor market. Increased product market competition may also involve improvements in production technologies (a larger A), such as ones brought about by the externalities associated with having a larger market. This may increase the laissez-faire equilibrium employment level with respect to its level before the shock, shifting the labor demand schedule upward. But under greater product market competition, the employment bias of labor market institutions with respect to the laissez-faire outcome is larger. To put it another way, if the rationale for labor market institutions is only in terms of (functional) income distribution, then the wedge should be downscaled after globalization because there is a steeper efficiency-equity trade-off.

Overall, an increase in product market competition leads to pressures to reduce the wedge that labor market institutions entail with respect to the competitive outcome. At the same time, however, unreformed labor markets have worse employment outcomes than before globalization. This means that stronger competitive pressures in product markets also increase the risk of job loss, potentially creating strong constituencies against the retrenchment of institutions that protect against unemployment risk, like nonemployment benefits, employment protection, and early retirement, whose reform pattern is characterized in table 1.1.

Minimum Wages

The minimum wage is a labor market institution that sets a wage floor, that is, a lower bound to the wage paid to individual workers. The first minimum wage was introduced in the United States in 1938 and paid 25 cents per hour. In 2007 the federal minimum wage was \$5.85, in nominal terms 23 times larger, but, in real terms, only 1.4 times larger than 70 years ago.

Although most countries in the world have some form of minimum wage, the scale, eligibility, and operational details change from country to country, so providing a cross-country-comparable definition and measure of the minimum wage is not an easy task. Some statistics, like the ratio of the minimum wage to the average wage, are, however, commonly used to summarize the relevance of the minimum wage in affecting the distribution of earnings in different countries. A key difference between the minimum wage and the other price-based institutions analyzed in the following chapters (e.g., unions and collective bargaining) is that the minimum wage mostly affects the low end of the wage distribution.

A large body of theoretical and empirical research examines the effects of the minimum wage. Theory offers clear-cut predictions only in the case of a competitive labor market. Empirical results point in both directions—positive and negative effects of the minimum wage on employment—which is possible in a labor market where individual firms face upward-sloping labor supply curves.

In some countries the minimum wage is unilaterally set by the government, while in other countries it is the outcome of negotiations between workers and firm representatives. When it is government legislated, the minimum wage in principle applies to all workers who have a legal contract. When it is the outcome of collective bargaining, the wage floor agreed to by the parties involved in wage negotiations may also cover the workers who are not unionized. In this case it becomes a minimum wage applied to all workers covered by the collective agreements. Overall, it is convenient to classify the different types of minimum wages applied in OECD countries according to their coverage and determination on the basis of the following threefold taxonomy:

1. A national, government-legislated (perhaps after consultations with trade unions and employers' associations) minimum wage

2. A national minimum wage that is the outcome of collective bargaining agreements and is extended to all workers
3. An industry-level minimum wage that results from industry-level collective bargaining and is extended to all workers in that industry

All these minima can be set on an hourly, daily, weekly, or monthly basis. Beyond the single minimum wage, there is often a reduced or subminimum rate for some specific groups of workers, for example, those without work experience and/or youngsters. Often, subminimum rates do not exist *de jure* but exist *de facto*, since special employment programs allow employers to pay lower wages to youth workers. In some countries premiums to the minimum wage are allowed to reward specific and time-varying workers' characteristics. For instance, the minimum wage may increase with workers' experience, workers' qualifications, and family status. Legislated minimum wages may or may not be automatically indexed to price inflation; in the latter case they may be discretionally adjusted by governments.

Minimum wages bear a close relationship to other institutions acting over the entire wage distribution. Minimum wages are integrated into union wage platforms (chapter 3) when the minimum wage is set and enforced under collective agreements. Active labor market policies (chapter 12), notably employment-conditional incentives, also interact with the minimum wage to reduce its potential disemployment effects.

2.1 Cross-Country Comparisons

Despite the various differences that we have pointed out, one can still try to compare minimum wages across countries by measuring their value relative to some central measure of the wage distribution. In particular, the ratio of the minimum wage to the average wage is often used in international comparisons. In principle, using the median rather than the average wage as a denominator would be preferable because the average wage may be affected by large values at the upper tail of the distribution, but micro data are not always available to obtain the median wage. This measure is clearly affected by how both the numerator (the minimum wage) and the denominator (the average wage) are measured. Because the minimum wage is generally exempted from income taxes, which are often progressive, it may also be preferable to use net wages (average wages after taxes) as the denominator in computing this measure. Particular attention should also be paid to using the appropriate earnings measure, which should possibly exclude any overtime and bonus payments.

The ratio of minimum to average wage ignores potential spillovers associated with the setting of minimum wages. Especially when minimum wages are embedded in collective bargaining, an increase of the minimum wage may also induce

pay increases above the minimum, shifting a relevant portion of the wage distribution to the right and hence leading to a significant increase of the average wage. Under these circumstances a change of the minimum wage will hardly be perceived from looking at the ratio of minimum to average wage, because both the numerator and the denominator move in the same direction.

Another drawback of this measure is that it does not take into account the fact that there may be subsets of the workforce, such as informal-sector workers, who are not covered by the minimum wage. In most developing countries and in a number of OECD countries (including formerly planned economies of central and Eastern Europe and southern European countries), alongside a formal (and often urban) job sector where the minimum wage is enforced, there is a large informal labor market in which the minimum wage legislation does not apply. There is no country in the world where there are enough labor inspectors to check every plant. Because of these enforcement issues, an increase of the minimum wage may paradoxically reduce the wage of the lowest-paid workers because low-productivity workers, crowded out from the covered sector by the rise in the minimum wage, could increase labor supply in the uncovered sector, driving down wages therein.¹

With these caveats in mind, table 2.1 displays the ratios of minimum to average wage for several countries as compiled by the OECD, taking as the denominator the wage of the average production worker (APW) for purposes of international comparison. As shown by the table, there is a wide range of values of the index, which goes from a low of 27 percent in Korea to a high of 52 percent in Ireland and Portugal. Minimum wages are lower in Japan, the United States, and Canada than in many European countries. Notice that the current cross-country asymmetries are the by-product of diverging historical developments: as indicated by figure 2.1, the minimum wage in the United States has been falling since the end of the 1960s relative to the average wage, just when it started steadily increasing in France.

The New Member States of the European Union (Hungary, Czech Republic, and Poland) are at the low end of the European distribution of minimum wages, together with Spain and UK, which introduced the minimum wage in 1999. Nominal values of monthly minimum wages in euros are also offered in the second column of the table, and in the third they are transformed into purchasing power units to make this number comparable across countries. Finally, the fourth and fifth columns provide information on the type of minimum wage (i.e., set by law or outcome of collective bargaining) and on its coverage.

¹ The *Kaitz index* (Kaitz 1970) was developed with the idea of taking into account the actual coverage of the minimum wage. It is defined as the ratio of the minimum wage to the average wage adjusted for the industry-level coverage of the legislation, but the eligible population is not always well defined.

Table 2.1 Minimum Wages in OECD Countries (2005)

	Minimum wage to average wage ratio ¹ (%)	Minimum wage (€ per hour)	Minimum wage ² (€ per month) PPP	Determination ³		Coverage ⁴
				Setting	Level	
Australia		7.25	1277	–	–	80
Austria				CB-L	P	95
Belgium	43	6.93	1220	CB	N	90
Canada	35	4.75	836	L	F-P	100
Czech Republic	39	1.58	278	L	N	100
Denmark				CB	–	80
Finland				CB	N	90
France	52	7.51	1322	L	N	100
Germany				CB	–	68
Greece		3.29	578	L	N	100
Hungary	38	1.28	225	L	N	100
Iceland				CB	–	–
Ireland	53	7.43	1308	CB	N	100
Italy				CB	N	80
Japan	40	4.15	731	L	P	100 ^a
Korea	27	2.64	464	–	–	10
Luxembourg				L	N	100 ^b
Netherlands	39	7.30	1284	L	N	100 ^c
New Zealand	48	4.98	877	L	N	25
Poland	40	1.35	237	L	N	100
Portugal	53	2.08	366	L	N	100
Slovak Republic				L	N	100
Spain	40	3.40	599	L	N	100
Turkey		2.78	489	L	–	100
United Kingdom	39	6.40	1127	L	N	100 ^d
United States	31	3.48	613	L	N	100

Sources: OECD, LFS Database; ILO Minimum Wage Database.

Notes: ¹ Minimum wage as percentage of the wage of an average production worker (APW).

² Real hourly minimum wage in Purchasing Power Parity times monthly number of hours.

³ Indicates whether wage floors are set by statutory rules defined by law or by collective negotiation and the levels of this agreement. CB = collective bargaining; L = set by law; P = provincial; F = federal; N = national.

⁴ Coverage is equal to 100 if the minimum wage is set by law, or where the coverage of collective agreements is extended to all workers; otherwise it measures the fraction of workers covered by the collective agreements defining contractual minima.

^a At the end of the 1990s.

^b For workers over 18 years old.

^c For workers over 23 years old.

^d For workers over 22 years old.

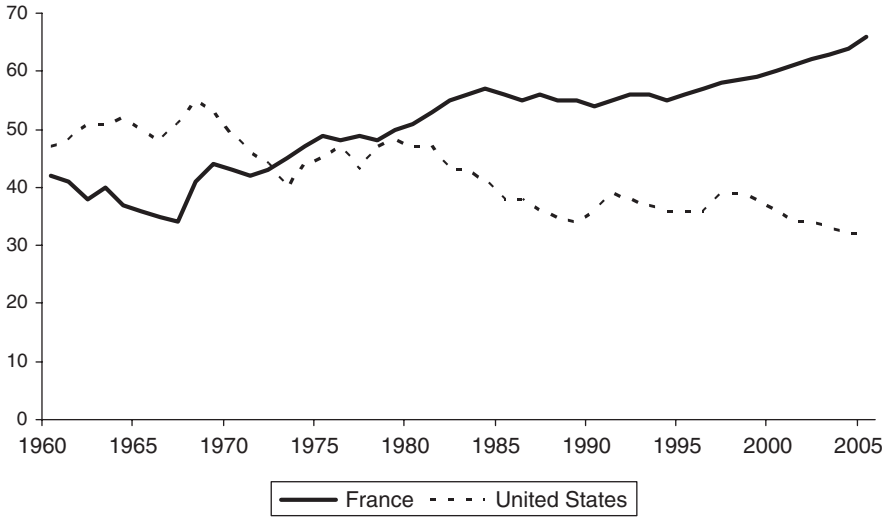


Figure 2.1 Ratio of Minimum to Median Wage

Source: OECD Minimum Wage Database.

2.2 Theory

2.2.1 A Competitive Labor Market

Economic theory offers unambiguous predictions about the effects of a minimum wage in a competitive labor market. A minimum wage set above the market-clearing level tends to reduce employment and increase the equilibrium wage level. Because the wage actually paid by employers is higher and employment is lower, some workers who were previously working at a lower wage are displaced by the introduction of the minimum wage, while other workers who were not supplying labor at the market-clearing level are now willing to work at the minimum wage. As a result of these two effects—displacement of some workers and higher participation—the introduction of a minimum wage above the market-clearing level involves some unemployment. As shown by figure 2.2, the minimum wage, $\underline{w}$, changes the slope of the labor supply schedule, preventing employer-firms from hiring workers at a lower wage than $\underline{w}$ even when their reservation wage is lower than the minimum wage. The presence of a minimum wage flattens the effective labor supply faced by employers, which becomes horizontal to the left of the locus $L^s(\underline{w})$. Notice further that the segment $L^s(\underline{w}) - L^d(\underline{w})$ denotes *unemployed* individuals, that is, persons who are not working, but who would be willing to supply labor at the minimum wage. Insofar as their reservation wage is lower than $\underline{w}$, these individuals will not be indifferent between working and not working: there is strictly a welfare loss associated with unemployment.

2.2.2 A Noncompetitive Labor Market

The effects of the introduction of a minimum wage in a labor market with distortions are much harder to predict. In some circumstances, for example, when employers have monopsony power in wage setting, notably when there are matching frictions and externalities associated with job search, the introduction of a minimum wage may actually end up increasing employment.

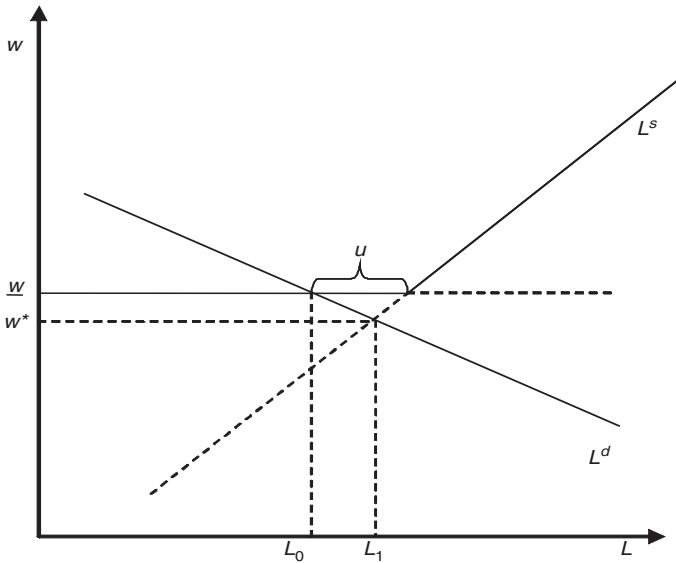


Figure 2.2 The Minimum Wage in a Competitive Labor Market

When employers can unilaterally set wages, their profit-maximizing choice involves lower employment and wage levels than in a competitive economy. As first pointed out by Stigler (1946) and Lester (1947), there is a theoretical possibility that a minimum wage set above the equilibrium wage could increase employment. The reason why this happens can be readily grasped by turning upside down the well-known case of a monopoly. As a pure monopolist (in the product market) faces a downward-sloping *demand* curve for its products, a pure monopsonist (in the labor market) faces an upward-sloping labor *supply* curve. This means that the marginal cost of hiring a worker for this unique employer is higher than the reservation wage of any additional worker, because the pay increase necessary to induce the individual to supply labor has to be granted not only to the marginal worker, but also to the entire workforce (just as an increase in the supply of a monopolist involves a decline in the price of all goods being sold, not just of the last unit of output). The marginal hiring cost (*mhc*) of a monopsonist is depicted in figure 2.3: it is above the $L^S(w)$ curve and deviates more and more from the latter

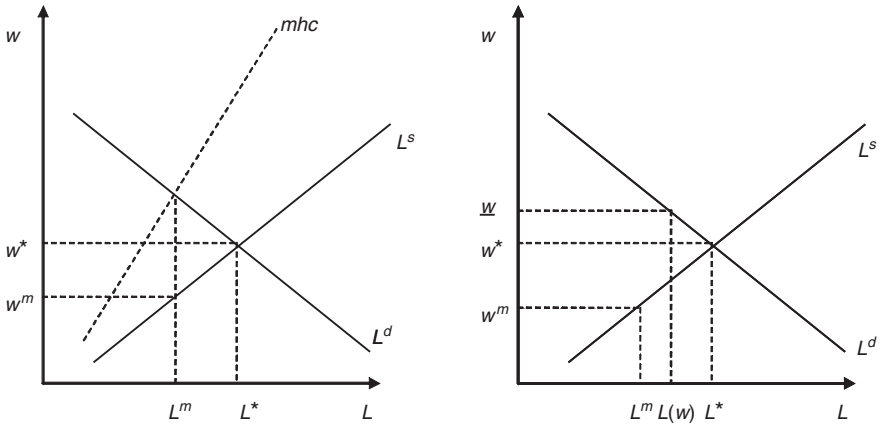


Figure 2.3 Monopsony and the Minimum Wage

as hiring an additional workers implies a wage increase to a larger and larger workforce. The profit-maximizing employment choice of a monopsonist then equals the marginal hiring cost to the marginal revenues from labor, $L^d(w)$ (the price of the good multiplied by the marginal productivity of labor). Graphically, this optimal choice lies at the intersection of the marginal hiring cost and labor demand curves. The monopsonist hires L_m workers, compared with L^* in a competitive economy. It also pays a lower wage (w^m) than any firm in a competitive equilibrium (w^*) because a monopsonist pays at the equilibrium less than the marginal productivity of labor. The degree of *monopsonistic power* of a firm is measured by this wedge between labor demand and supply, $y(L_m) - w^m$. This wedge is larger the steeper the labor supply schedule. More precisely, it is decreasing with the responsiveness of labor supply to wages, as shown analytically in box 2.1.

Box 2.1 The Degree of Monopsonistic Power

The results in figure 2.3 clearly have an analytical counterpart. Denote, as usual, by $y(L)$ the value of the marginal product of labor and by $G(w)$ the aggregate labor supply faced by the monopolist. At the monopsony equilibrium the value of the marginal product of labor must equal the marginal cost of labor:

$$y(L_m) = w^m(1 + \varepsilon), \quad (2.1)$$

where ε denotes the inverse wage elasticity of labor supply. The wedge between labor demand (the value of the marginal product of labor for the firm) and supply (the way in which labor is rewarded) expressed as a fraction of the latter measures the degree of monopsonistic power of the firm. By rearranging (2.1) it is easy to

(continued)

Box 2.1 (continued)

derive an expression for the degree of monopsonistic power,

$$\frac{y(L_m) - w^m}{w^m} = \varepsilon,$$

which tends to zero as the wage elasticity of labor supply tends to infinity (as ε tends to zero).

Thus monopsonistic power is decreasing with the wage elasticity of labor supply: the less elastic the labor supply, the larger the difference between the equilibrium value of a job for the monopsonist and the wage that the monopsonist pays to workers. Conversely, when labor supply is infinitely elastic (the employer is a price-taker in the labor market), monopsonistic power is zero.

In the presence of a monopsony, a minimum wage set at an intermediate level between the monopsony and the competitive economy levels increases both wages *and* employment. However, if the minimum wage happens to be above w^* , as depicted in figure 2.3, then the minimum wage involves instead a lower employment level than at the competitive economy equilibrium. For even higher values of the minimum wage, employment falls below the level attained in the pure monopsony equilibrium. In the presence of a monopsony, there is therefore a nonmonotonic relationship between the minimum wage and employment: for sufficiently low levels of the minimum wage, an increase in the minimum wage is accompanied by an increase in employment, while above some threshold the traditional negative relationship exists.

The case of a minimum wage in the presence of a pure monopsony is often neglected because it is considered of limited practical relevance. Labor markets with only one employer are rarely observed. The standard example is the one-company towns in Russia, a legacy of central planning. Like mountain gorillas, these monopsonies are a sort of endangered species. A case like that described by figure 2.3, however, may also arise when there is more than one employer in a labor market, but these employers collude in wage setting. Collusion among employers may be favored by collective bargaining institutions, but collective bargaining means that also the worker side is organized, and hence employers cannot unilaterally set wages. There are in such cases *bilateral monopoly* conditions that reduce the power of employers in setting wages.

Even labor markets in which each individual employer is infinitesimally small relative to the market as a whole, however, may confer on employers some degree of monopsony power. This happens when there are search frictions and mobility costs, making it costly for workers to change jobs. These costs prevent the labor market from arbitraging away any difference in the way in which different employers pay for (homogeneous) labor services. Job creation and hiring costs

also attribute to employers some degree of monopsony power. When it is costly to establish a new job or hiring is highly regulated, there are fewer vacancies, and hence for workers it is harder to find openings to apply for. All these modern monopsony cases are rather frequent in practice (Manning 2003) and arise even when there are many employers, but many fewer vacancies to apply for.

Can a minimum wage increase employment also in the presence of these modern monopsonies? Yes, if the minimum wage is not too large and workers' search efforts and participation decisions are sufficiently responsive to the wage that they can earn if they find a vacancy. The intuition is straightforward. In a labor market without frictions, a minimum wage, on the one hand, reduces labor demand (we move up and to the left along the L^d schedule in figure 2.2), but, on the other hand, it also increases labor supply (we move up and to the right along the L^s schedule). In a labor market with frictions and where individuals can choose whether or not to search, as well as the level of search intensity, a higher wage paid to those who succeed in finding a job induces more people to participate in the labor market and put effort into job search. At the same time, the minimum wage reduces the surplus that employers can earn over and above the marginal productivity of workers once a vacancy is filled. However, the presence of more jobseekers and the fact that each of them is seeking jobs more intensively increases the probability that a vacancy is filled. To the extent that this second, labor supply effect dominates the negative labor demand effect associated with the reduction in the surplus of employers, there will be both more jobseekers and more vacancies at the equilibrium with the minimum wage than at the equilibrium without the minimum wage. Because the labor market is larger, equilibrium employment will also be larger, even if each individual worker now faces a lower job-finding rate per any given level of search effort. Minimum wages can also have positive effects on employment when employers imperfectly monitor the productivity of workers in large firms. A minimum wage, in this context, may force firms to grow larger (Rebitzer and Taylor 1995). The intuition is that firms pay higher wages but also obtain higher productivity in return, as the penalty associated with disciplinary layoffs is larger.

Minimum wages may not have negative effects on employment in *dual labor markets* where the minimum wage does not apply to the secondary or informal labor market. Under these conditions there are important spillover effects between the two sectors. As pointed out by Gramlich (1976), Mincer (1976) and, Welch (1976), after a minimum wage increase, workers displaced in the formal sector move to the uncovered sector. Hence, as depicted in figure 2.4, wages in the informal sector fall (from w_o^I to w_1^I), and labor supply in the formal sector declines (shifting the L^s curve to the left). The minimum wage then reallocates jobs from the formal to the informal sector, increasing the difference between formal and informal wages. This adjustment mechanism prevents employment losses only insofar as there is perfect labor mobility between the two sectors.

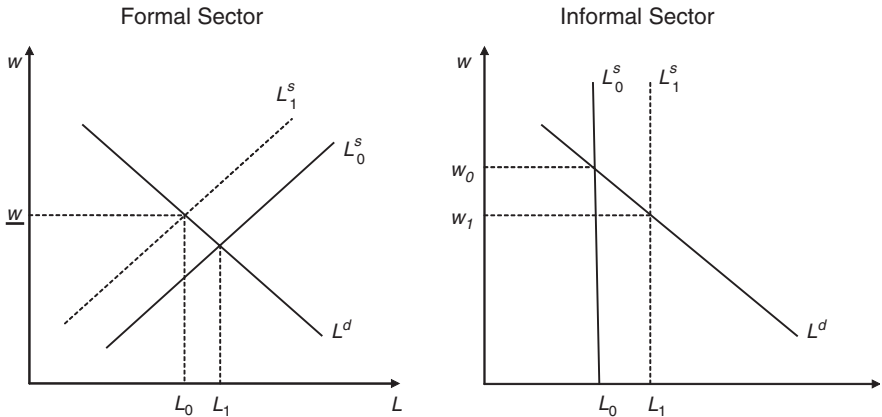


Figure 2.4 A Dual Labor Market and the Minimum Wage

Finally, positive effects of the minimum wage on welfare, although not necessarily on employment, can be generated when the productivity of a job depends on the investment in education and training by the employee (Cahuc and Michel 1996; Acemoglu and Pischke 1999). Under these conditions a binding minimum wage induces workers to raise their productivity by acquiring education in order not to be crowded out by the minimum wage. Hence, at the equilibrium with the minimum wage, there are more high-productivity jobs. The same type of effect may be induced on the demand side (Acemoglu 2001) because the minimum wage increases the number of vacancies for high-productivity jobs issued by employers.

Overall, although the standard prediction from economic theory is that a minimum wage should reduce employment, a number of market imperfections may allow the introduction of a minimum wage, set at relatively low levels, to be consistent with the attainment of higher levels of employment and welfare.

2.3 Empirical Evidence

2.3.1 Studies Based on Firm-Level Data

Many studies of the effects of the minimum wage are based on firm-level data and estimate the impact of minimum wages on labor demand. The impact of the minimum wage is clearly dependent on the characteristics of the bottom end of the wage distribution and on the actual enforcement of minimum wages. Earlier studies used the Kaitz index, to control for limited enforcement of the minimum wage. More recent empirical work typically measures the proportion of people earning a wage between the old and the new minimum wages (Card 1992; Card and Krueger 1995; Brown 1999), or the *fraction affected* by the minimum wage

increase. An increasingly used measure of enforcement is the *spike*, defined as the proportion of people earning exactly one minimum wage (Dolado et al. 1996). If the minimum wage is properly enforced, we generally expect the spike to increase after a minimum wage is raised as lower wages are aligned to the new minimum.

Dolado et al. (1996) provide an overview of studies of the effects of the minimum wage on employment in several OECD countries. The minimum wage is generally found to affect employment negatively, although the magnitude of the effects varies from country to country and depending on the category affected (e.g., minimum wages for youngsters generally have stronger negative effects on employment). OECD (1998, 2006c) more recently reviewed the empirical literature on the effects of the minimum wage and found some, albeit small, negative effects on employment of the minimum wage, notably for young workers.

Although most of these studies use data on the formal sector, where minimum wage legislation is enforced, a few studies apply the same methodology to data on the informal sector. These studies are mostly concentrated on developing countries (Lemos 2004 for Brazil, Gindling and Terrell 2004 for Costa Rica, and Jones 1997 for Ghana), where the informal sector is larger. Notwithstanding problems in measuring informal-sector employment, a few studies surprisingly found an increase in wages also in the informal sector after a minimum wage hike. The interpretation provided by this literature is that the minimum wage of the formal sector serves as a reference throughout the economy. If firms have monopsonistic power in the informal sector, and fair remuneration considerations are relevant, it is possible that changes in the minimum wage in the formal (and covered) sector lead to corresponding increases in the average wage of the informal sector. The term *lighthouse effect* has been used to denote this phenomenon.

2.3.2 Studies Based on Natural Experiments

Most of the studies just reviewed compare employment and wage outcomes of workers whose wage has to be raised to comply with the minimum wage with employment and wage outcomes of workers higher up the wage distribution, presumably unaffected by changes in the minimum wage. The problem with this approach is that persons receiving the minimum wage are not representative of the entire population. Thus we may end up attributing to the minimum wage effects that are related to different characteristics of workers (e.g., lower labor market attachment) located at varying portions of the wage distribution.

The empirical methodology of a natural experiment in economics (Meyer 1995; Angrist and Krueger 1999; Blundell and Dias 2000) takes these selection problems into account and makes it possible to better identify the effects of the minimum wage. It consists of exploiting exogenous changes in the economic environment

of certain agents in order to compare their reactions to those of other (a priori identical) agents who have not undergone these changes. In practice this means finding a counterfactual or *another difference* that makes it possible to control for the difference in the composition of those affected by the minimum wage and those not affected by it. For this reason, the estimators obtained by applying this methodology are also called *difference-in-differences* or double-difference estimators.

Difference-in-differences estimators have been used, for instance, by Stewart (2004) to investigate the effects of the introduction of the National Minimum Wage in Britain in 1999 (with the adult rate set at 3.6 pounds per hour) and of the subsequent increases in 2000 (3.7) and 2001 (4.1). As documented by table 2.1, Britain has a ratio of minimum wage to median wage located roughly in the middle of the distribution of European countries. Stewart compared employment outcomes of individuals just above the minimum wage and higher up the wage distribution before and after the introduction of the minimum wage. He repeated this exercise for different demographic groups (males and females at different age groups) and controlled for cyclical conditions. Stewart found no adverse effect of the introduction of the minimum wage in Britain for any of the demographic groups considered.

Difference-in-differences estimators have also been extensively used in the United States, exploiting cross-state variation in setting minimum wages above the federal level. Card and Krueger (1994, 1995), in particular, investigated the impact of increases in the minimum wage in New Jersey in 1992 from \$4.25 to \$5.05. They used as the control group Pennsylvania, where the minimum wage remained at \$4.25 throughout this period. New Jersey and Pennsylvania are bordering states and have similar economic structures. Because this study has been widely debated, it is discussed in some detail in box 2.2.

Box 2.2 Effects of Minimum Wage Hikes in the U.S. Fast-Food Industry

David Card and Alan Krueger collected data on employment in 410 fast-food restaurants in New Jersey and Pennsylvania, two bordering states of the United States with similar economic structures. The minimum wage was initially the same (\$4.25 per hour) in both states and was raised in 1992, only in New Jersey, to \$5.05 per hour. The data were collected in March 1992 (when both states had the same minimum wage) and in December 1992 (after the increase of the minimum wage in New Jersey). The changes in the wage distribution in the two U.S. states are totally characterized in figure 2.5: there is a visible shift of the distribution in New Jersey, with a spike corresponding to the new level of the minimum wage, while the wage distribution in Pennsylvania remains roughly unchanged.

(continued)

Box 2.2 (continued)

The effects of the minimum wage increase on employment were estimated by Card and Krueger by simply taking the difference between the December 1992 and March 1992 employment variations in the two states. The rationale for this method is as follows. Suppose that employment (L) in any state is determined by an equation of the type (we omit error terms for simplicity)

$$L_i = \alpha w_i + X_i \gamma,$$

where w_i is the level of the minimum wage (the policy) in state i and X_i contains all the other variables that influence L_i . If we have two observations that refer to dates before and after the policy reform for the state changing the minimum wage (New Jersey), we can obtain

$$\Delta L_i = L_{i2} - L_{i1} = \alpha(w_{i2} - w_{i1}) + (X_{i2} - X_{i1})\gamma,$$

where the second subscript indicates dates before (1) and after (2) the change of the minimum wage. In words, the total variation in employment in the treatment state is the sum of two components: (1) the effect of the policy change itself (the increase in the minimum wage), and (2) the effect of changes in the other variables affecting employment. In order to isolate the effect of the policy (in order to estimate α), we need to find another difference, for example, get time-series data for another state j that is identical to i in each characteristic except that it has not carried out the reform (that is, $w_{j2} - w_{j1} = 0$). This is the role played by Pennsylvania in Card and Krueger's study. The variation of employment in state j is given by

$$\Delta L_j = L_{j2} - L_{j1} = \alpha(w_{j2} - w_{j1}) + (X_{j2} - X_{j1})\gamma = (X_{j2} - X_{j1})\gamma.$$

As long as the two states are sufficiently similar ($X_i = X_j$ at both dates, 1 and 2), we can obtain an estimate of α by simply calculating the difference of the difference:

$$\Delta L_i - \Delta L_j = \alpha(w_{i2} - w_{i1}).$$

In the case of Card and Krueger, the variation of the minimum wage ($w_{i2} - w_{i1}$) in New Jersey was 80 cents. They found the difference of the differences to be

$$\Delta L_i - \Delta L_j = 0.29 - (-2.01) = 2.30;$$

hence they obtained

$$\hat{\alpha} = \frac{\Delta L_i - \Delta L_j}{(w_{i2} - w_{i1})} = \frac{2.30}{.8} = 2.875,$$

which implies that an increase of the minimum wage by one dollar has the potential to create 2.875 more employees.

(continued)

Box 2.2 (continued)

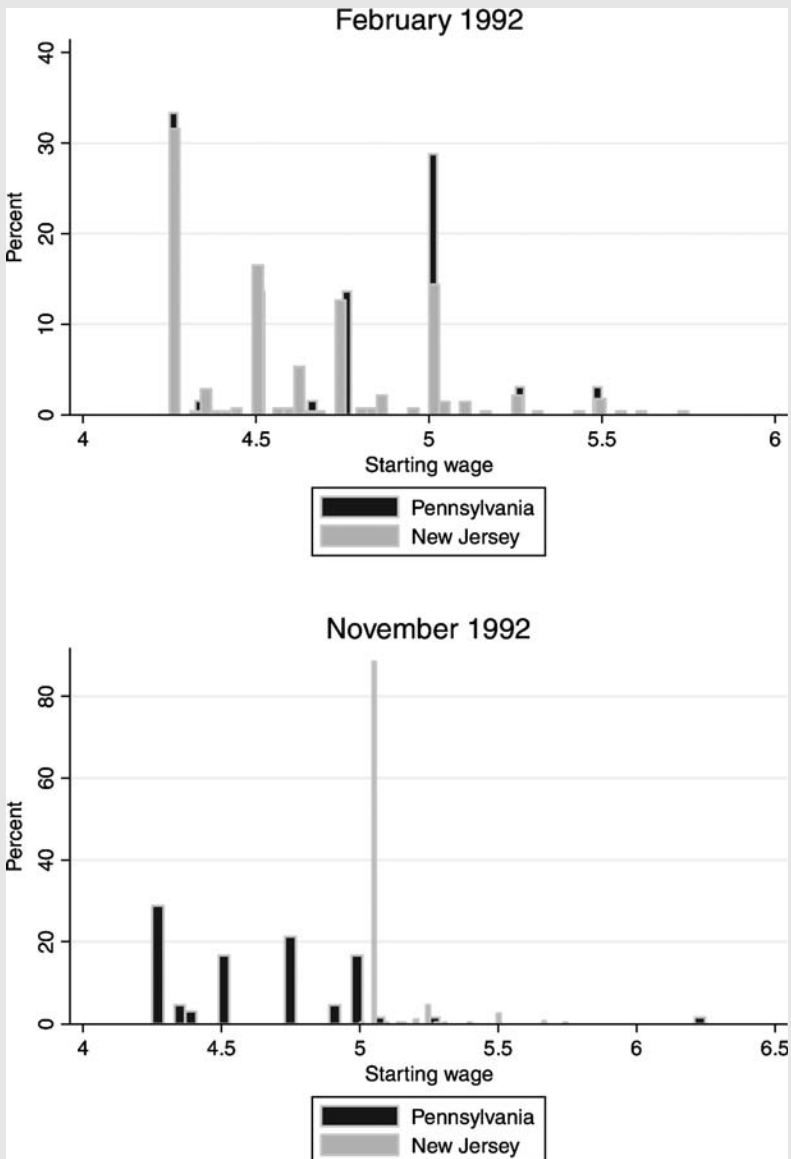


Figure 2.5 The Wage Distribution before and after the Increase in the Minimum Wage

(continued)

Box 2.2 (continued)

Hence Card and Krueger found that after the minimum wage was increased, the level of employment in fast-food establishments in New Jersey rose faster than in Pennsylvania. Their conclusion was that therefore an increase in the minimum wage can lead to an increase in employment when this wage is sufficiently low to start with.

The research by Card and Krueger generated a large and very informative debate along two dimensions. The first debate focused on the interpretation of the results and on whether consumers of fast food can be taken to be representative of the population at large, since it may well be that persons earning minimum wages represent the typical consumers of fast food. The second debate concerned the quality of the data and the fact that the original Card and Krueger study was based on a survey that used telephone interviews and was not the result of administrative data. Despite the large debate between Neumark and Wascher (2000) and Card and Krueger (1994), the results of the early study seem to be confirmed. A Stata data file with the Card and Krueger dataset and a do file and a log file are available on our webpage (see link at <http://press.princeton.edu/titles/8771.html>).

Sources: Card and Krueger (1994); Neumark and Wascher (2000).

2.3.3 Studies Based on Workers' Histories

Two problems with this literature are that (1) it typically focuses on specific industries in analyzing the effects of the minimum wage, while the standard predictions of the competitive model apply to the labor market as a whole, and (2) it neglects potential effects of the minimum wage on hours rather than on persons employed.

These issues can be tackled by using longitudinal data on representative samples of workers, tracking labor market histories of persons whose wages are at the minimum wage or close to it. Recent studies in this category found that changes in the minimum wage have a significant impact on employment among this group of workers. Nevertheless, there is no agreement about the directions of these changes. Abowd et al. (1999) found that in France an increase of 1 percent in the minimum wage reduces the probability that men receiving the minimum wage keep their previous jobs by 1.3 percent, while for women this figure is 1 percent. In the United States a reduction by 1 percent in the minimum wage increases the probability that workers paid at this level will keep their jobs by 0.4 percent for men and by 1.6 percent for women. Portugal and Cardoso (2001) found different results using the same type of methodology. They exploited changes made in 1987 to Portuguese legislation regarding the minimum wage of young people aged 19 and under. The minimum wage was raised by 50 percent for youths aged 17 and by 33 percent for youths aged 18 and 19. They found that these minimum wage hikes reduced hiring, but also that workers had a greater tendency to keep their jobs. In other words,

Portugal and Cardoso observed fewer separations, which partly offset the fall in hires. Note that this result is consistent with the predictions of the monopsony model, since it reveals greater attachment of youth to their jobs when wages improve. Neumark et al. (2004) found negative effects of minimum wage hikes in the United States on both employment and hours of workers initially earning the minimum wage or slightly more, contrary to Zavodny (2000), who found that an increase in the minimum wage reduced the probability that an affected worker remained employed, but for those who kept their job, there was a positive effect on hours.

Overall, this large body of empirical research shows that the minimum wage can have significant effects on both job-finding and job loss probabilities. However, it does not invariably find a positive effect on the probability of job loss among the affected population.

2.4 Policy Issues

Thus recent empirical evidence fails to provide unambiguous results on the effects of the minimum wage on employment.² In particular, only two-thirds of the studies reviewed by Neumark and Wascher (2007) found negative employment effects of minimum wages, and these effects were not always statistically significant. This explains why some researchers advocate an increase of the minimum wage, while others argue in favor of marked reductions of the minimum wage.

2.4.1 Should the Minimum Wage Be Reduced or Increased?

Models of the minimum wage under realistic conditions, that is, allowing for some degree of monopsony power by individual firms, suggest that the setting of the minimum wage is a matter of fine-tuning: if it is too low, it is not binding; if it is too high, it can do worse than the market failure that it was supposed to address, where “worse” here means that the total surplus is lower than without the minimum wage.

The main rationale for advocating a reduction in the minimum wage is that the labor market is ill functioning, and some low-productivity workers (e.g., youngsters and the unskilled) inflate the ranks of unemployment. This argument is clearly stronger in the presence of double-digit unemployment rates for these groups, notably when these rates are not paralleled by adverse labor market conditions for other workers.

An increase in the minimum wage is often advocated on the grounds that some groups of workers have a particularly weak position at the bargaining table, and

² As Flinn (2007) puts it, “Recent studies indicate that the ‘textbook-competitive’ model of the labor market ... may have serious deficiencies in accounting for minimum wage effects on labor market outcomes.”

levels of earnings inequality are deemed to be too large. In this context the minimum wage is seen as an instrument to reduce the number of working poor, that is, individuals who hold a full-time job but nevertheless appear to live close to the poverty line. The natural supporters of this view are those who are likely to enjoy a wage increase in the presence of a wage floor.

2.4.2 Is the Minimum Wage Effective in Reducing Earnings Inequality and Poverty?

As is apparent from the preceding, the strongest arguments in favor of an increase in the minimum wage rely on equity considerations. However, economic theory also does not provide firm guidance on the impact of the minimum wage on poverty. A working poor person employed at the minimum wage may experience an increase in income if that person's job is not destroyed, reducing the poverty rate (the percentage of individuals having incomes below the poverty line), but if the minimum wage hike destroys jobs, some individuals will experience a drop in their incomes, increasing the incidence of poverty (the distance between the average incomes of those above and below the poverty line), if not the poverty rate itself (Brown 1999). In dual labor markets an increase in the minimum wage could quite paradoxically end up increasing earnings inequality. Some of the adverse effects of minimum wages on unemployment among low-productivity workers can be mitigated when the minimum wage is combined with in-work benefits (chapter 4). Actually, the combination of minimum wages and in-work benefits is often advocated as a rather effective antipoverty device (Gregg 2000; OECD 2006c), providing *wage insurance* to those with low earnings.

Moreover, a substantial portion of minimum wage earners may not be poor because other family members have earnings. Thus the minimum wage may have a low *target efficiency*, helping many workers in nonpoor families and providing only limited earnings support to the truly needy.

Because economic theory does not offer unambiguous results about the effects of the minimum wage on poverty, it is always important to examine this issue empirically. Applied studies typically look at the distributional impact of the minimum wage by analyzing wage distributions in a neighborhood of the minimum wage level. If crowding-out effects are important, then we should observe a spike in the wage distribution close to the legally imposed minimum. If there is no spike or the spike lies to the right of the minimum, the data indicate that the minimum wage has little effective "bite." Most studies (including Card and Krueger 1994; see figure 2.5 in box 2.2) actually found a spike in the wage distribution corresponding to the minimum wage. There is also less ambiguity in the empirical literature on the employment effects of minimum wages when the focus is on teenage and unskilled workers: in this case the evidence

of disemployment effects is particularly strong. Concerning the effect of the minimum wage specifically on poverty, the study of Addison and Blackburn (1999) suggests that the increases in the minimum wage that occurred in the United States in the 1990s contributed to reducing the poverty rate among youth aged 24 and under and workers over 24, but only among school dropouts. Finally, Flinn (2002) found positive effects on welfare of young Americans for the 1997 minimum wage increase (from \$4.75 to \$5.15), but no evidence of a positive effect on welfare of the minimum wage increase that occurred in 1996 (from \$4.25 to \$4.75), which seems not to have exerted significant spillovers on the wage distribution.

2.5 Why Does a Minimum Wage Exist?

Minimum wages can achieve both goals typically assigned to labor market institutions:

1. They can increase efficiency by remedying market failures, such as those deriving from excessive monopsonistic power, and informational asymmetries that give rise to moral hazard and adverse selection problems.
2. They can reduce earnings inequality by supporting incomes of low-earning, workers, for example, low-skilled workers.

The setting of the minimum wage requires careful fine-tuning if either of these two goals is to be achieved. When the minimum wage is too low, it is ineffective. When the minimum wage is set at a level that is too high, it reduces welfare and may have perverse effects on income inequality by completely crowding low-skilled workers out of the labor market.

In order to improve the efficiency and distributional properties of a minimum wage, governments have to adjust it over time, but it may prove politically difficult to do so. The political economy literature on the minimum wage (Sobel 1999; Saint-Paul 2000; Bacache-Beauvallet and Lehmann 2005) highlights which institutional features can increase or reduce political support for the minimum wage. The decisive (median) voter is generally an employed worker whose wage is slightly above the minimum wage. The key dimension along which to assess political support for the minimum wage is whether this pivotal group of workers experiences (1) a sufficiently high degree of substitutability with workers who are potentially crowded out by the minimum wage and (2) a positive wage spillover from the introduction of the minimum wage or a high degree of complementarity with capital. When an increase in the minimum wage, eliminating the least skilled, increases the marginal value of the semiskilled and hence their wages, the ruling middle class will support a rather high minimum wage. Conversely, the

ruling middle class will oppose a reduction in the minimum wage when it fears that firms will try to replace them with cheaper workers.

Hence the future of the minimum wage is likely to depend on these cross-skill (and capital-labor) complementarities and on the spillovers that the minimum wage can exert over the entire wage distribution. As discussed in chapter 3, these spillover effects are likely to be more pronounced in the presence of strong unions, setting wage scales from this minimum.

2.6 Suggestions for Further Reading

A good starting point is the book by David Card and Alan Krueger (1995) summarizing also their controversial study, which is described in some detail in box 2.2 and can be replicated with the data provided on the website (<http://press.princeton.edu/titles/8771.html>). The debate over their seminal article (Card and Krueger 1994) is also particularly instructive: in particular, we recommend David Neumark and William Wascher (2000) and the reply by Card and Krueger (2000) in the same issue of the *American Economic Review*. For a survey of the new minimum wage research, see Neumark and Wascher (2007). The earlier literature, mostly focused on the United States, is surveyed by John Kennan (1995).

A special issue of the *Economic Journal* (114) was devoted in 2004 to the introduction of the minimum wage in Britain. Finally, Juan Dolado et al. (1996) offer a less recent but particularly instructive discussion of the fallacies of the common wisdom on the effects of the minimum wage that draws mainly on the European experience.

2.7 Review Questions

1. What is the Kaitz index? What are the pros and cons of this measure of the minimum wage?
2. Why are there so few workers earning the minimum wage?
3. Why are minimum wages typically age dependent?
4. When does a minimum wage increase employment?
5. Why, in your view, has an increase in the minimum wage been found to increase wages also in the uncovered (informal) sector?
6. When does a minimum wage increase welfare, although not necessarily employment?
7. How does a minimum wage affect poverty?
8. Who supports the presence of the minimum wage?
9. Why did Card and Krueger study the fast-food industry?

2.8 Technical Annex: Fine-Tuning of the Minimum Wage

Specify labor demand and supply as in technical annex 1.4. Suppose that labor demand is originated by just one employer facing the aggregate labor supply. This pure monopsonist (superscript m) chooses the employment level that maximizes profits:

$$\pi^m = \frac{AL^{1-\eta}}{1-\eta} - wL,$$

subject to being on the labor supply curve $L = w^{\frac{1}{\varepsilon}}$. Deriving the first-order condition and solving for wages, one obtains

$$w^m = \left[\frac{A}{1+\varepsilon} \right]^{\frac{\varepsilon}{\varepsilon+\eta}} < A^{\frac{\varepsilon}{\varepsilon+\eta}} = w^*; \quad (2.2)$$

hence the equilibrium with a monopsonist involves a lower wage than the equilibrium in a competitive labor market maximizing the total surplus (w^*). Substituting the monopsonist wage in the labor supply, we obtain the equilibrium employment level under a pure monopsony,

$$L^m = \left[\frac{A}{1+\varepsilon} \right]^{\frac{1}{\varepsilon+\eta}} < A^{\frac{1}{\varepsilon+\eta}} = L^*.$$

Hence employment is also lower than in a competitive labor market (L^*). Thus there is an efficiency loss associated with the presence of a monopsony because the economy achieves a lower total surplus than in a perfect labor market.

From this equilibrium, a minimum wage can contribute to reducing the dead-weight loss associated with the monopsony. In particular, any minimum wage that forces the monopsonist to pay at least $\underline{w}$ will increase employment, provided that the minimum wage is lower than the wage maximizing the total surplus,

$$\left[\frac{A}{1+\varepsilon} \right]^{\frac{\varepsilon}{\varepsilon+\eta}} < \underline{w} < A^{\frac{\varepsilon}{\varepsilon+\eta}}.$$

A minimum wage set in this range has an efficiency-enhancing role. When the minimum is larger than $A^{\frac{\varepsilon}{\varepsilon+\eta}}$, it is itself a source of inefficiency potentially leading to an even lower total surplus than under a monopsony. There will be a nonmonotonic relation between the minimum wage and employment: first increasing and then decreasing. Still, distributional considerations may dictate a minimum wage above the perfect labor market case.

Consider the outcome of the normative, social planning problem characterized in technical annex 1.4. We know that the socially optimal wage will deviate from $A^{\frac{\varepsilon}{\varepsilon+\eta}}$ by a factor μ that is a function of labor demand and supply elasticities, as

well as distributional weights of employers and workers. Suppose that society cares only about the surplus of employers; that is, $\beta = 0$. In this case the optimal $\mu = \frac{1}{1+\varepsilon} < 1$; that is, the wage at the social optimum is lower than the wage in a perfect economy. More precisely, the wage satisfying the social welfare maximization in this case is

$$\mu(\beta = 0)w^* = \left[\frac{A}{1 + \varepsilon} \right]^{\frac{\varepsilon}{\varepsilon + \eta}} = w^m;$$

that is, the social optimum coincides with the equilibrium under a pure monopsony. Intuitively, the planner maximizes the surplus of employers only.

Consider now the other extreme case, that is, when $\beta = 1$ and the planner cares only about the surplus of workers. Under these conditions $\mu = \frac{1}{1-\eta} > 1$; that is, the social optimum involves a wage above the competitive labor market case. This wage can be attained by setting a minimum wage above $A^{\frac{\varepsilon}{\varepsilon + \eta}}$. The minimum wage is the simplest way to introduce a wedge between labor supply and demand that alters the equilibrium from the laissez-faire outcome.

This page intentionally left blank

Unions and Collective Bargaining

Labor (or trade) unions are voluntary membership organizations. Like political parties and churches, they represent the interests of their members, and their nature is defined by their membership.

Historically, unions emerged in the eighteenth century in the United Kingdom and the United States as craft, occupational organizations providing mutual insurance to their members against unemployment, death, and sometimes old age. Later, well into the nineteenth century, they gradually became industrial unions representing workers in semiskilled positions, harder to replace with unemployed workers or immigrants. Increasingly at the beginning of the twentieth century they became national organizations aiming at representing all workers and having a stronger political role. Membership in these general unions was stronger among manual workers, while skilled workers were less keen to join unions because of their egalitarian wage platforms; nowadays unions are often stronger in the public sector.

The potential membership of unions also depends on the underlying *collective bargaining* system. In some countries bargaining takes place at the national level, involving unions and employers' associations, with the government acting as mediator. In other countries there are only sectoral wage agreements or even more decentralized, plant-level agreements. Hybrid or multilevel bargaining structures with both centralized and plant-level agreements also exist, notably in Europe. Depending on the structure of the bargaining system, industrial or craft unions coordinate themselves within national federations of unions. Sometimes unions also coordinate across countries, as in the case of the European Trade Unions Confederation (ETUC).

Because of the interests of their members and their ideological roots, unions tend to pursue egalitarian wage policies, reducing wage differentials by education and skill level. At the same time, what unions do affects their membership. Unions are generally not very successful in recruiting highly skilled workers and seem to be aging more rapidly than the workforce, which points to growing difficulties in involving young people in the activities of unions.

Trade unions typically bargain over all aspects of an employment contract: wages, working hours, overtime pay, fringe benefits, employment security, and

health and safety standards. They negotiate with employers on a collective basis, overruling (when incompatible with the collective agreement) or complementing individual contracts. By coordinating wage claims of a plurality of workers, unions force employers to pay for labor services at a rate above the reservation wages of otherwise uncoordinated individuals. The bargaining power of unions is related to the extent of wage coordination that unions achieve, and to the coverage of the collective contracts signed by the unions. National labor unions bargain with governments over the minimum wage (chapter 2) and labor laws, like employment protection legislation (chapter 10), the age of retirement (chapter 6), family policies (chapter 7), and unemployment benefits (chapter 11). Thus the activity of unions interacts with the influence of many other institutions, both price based and quantity based.

3.1 Measures and Cross-Country Comparisons

Unfortunately, not much effort has been spent to date on developing cross-country-comparable measures of the influence of trade unions. This is because unions are reluctant to provide reliable figures on their membership, and surveys may not be accurate in eliciting membership because workers may not be keen to disclose this information. Available data suggest that the presence and influence of trade unions vary considerably across countries, and have changed quite dramatically over time. The most visible change has been an increasing divergence between unions' *presence* at the workplace, as measured by the number of active members in the workforce, and unions' *influence*, proxied by the fraction of workers involved in agreements negotiated by unions. The difference between union influence and union presence, which is generally referred to as the *excess coverage* of unions, has been increasing over time.

3.1.1 Union Density

The most intuitive measure of the relevance of trade unions is given by the *union density* (or *membership*) *rate*, that is, the fraction of workers registered with some trade union. Unions often have nonworking individuals among their members. For instance, in Italy there are more pensioners than workers in the ranks of the three largest trade unions, CGIL, CISL, and UIL. Figure 3.1 plots union density rates (counting only members of working age) in the OECD countries for which data are available since the 1960s. The continuous line displays the weighted average of national union density rates, where weights are provided by the working-age population. The two dotted lines represent one (weighted) standard deviation above and below the mean, respectively. After reaching a peak in 1978, membership sharply declined, although not by the same degree in all countries and regions (the cross-country variation in membership rates widened over time, as indicated

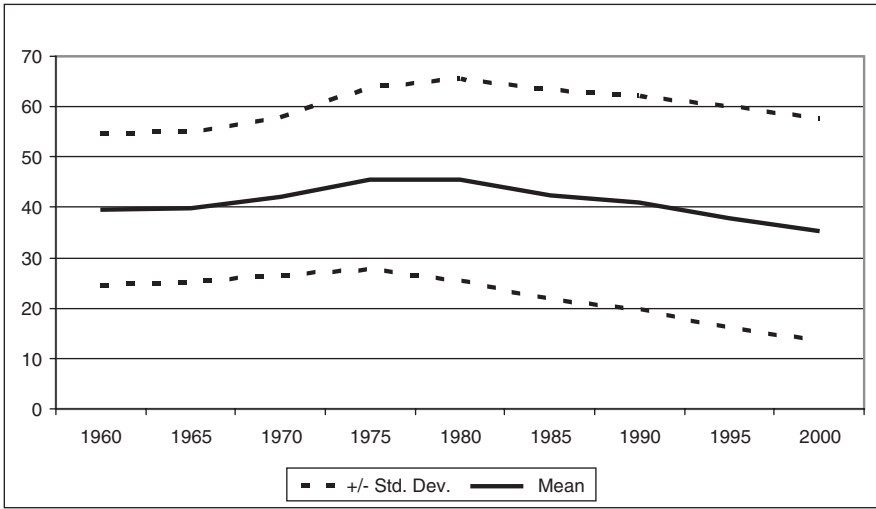


Figure 3.1 Union Membership in OECD Countries

Sources: OECD (2004) Ebbinghaus and Visser (2000).

Note: Weighted average of national union density rates, where union density is defined as members of working age as a fraction of employment and weights are the population shares.

by the larger band around the OECD average). Deunionization was particularly strong in the United States, where union density halved and nowadays membership in the private sector is below 10 percent, and the United Kingdom, where unions lost almost 4 million members in the 1979–87 period, during the Thatcher era. Strong declines in membership are also visible in Australia (from 48 to 23 percent), New Zealand (from 56 to 13 percent in the 1980–2003 period), and southern Europe.

The decline in union membership has gone hand in hand with the aging of unions. According to the Eurobarometer survey, in the period 1988–2001 the median age of a union member increased by more than two years in the four largest nations of the continental EU, France, Germany, Italy, and Spain, while, because of a significant reduction of youth unemployment, the age of a median worker was declining.

Membership, however, has not meant the same thing over time. Historically, membership in a trade union required mutual support, attendance at meetings, and involvement of time and effort in the running of the union’s internal affairs. Nowadays membership may require only voting over the Internet and paying a low voluntary contribution. In October 2005, for instance, the AFL-CIO enrolled members over its website (www.workingamerica.org), asking for \$5 voluntary dues. The reduction in union dues should encourage membership. At the same time, unions no longer or not always provide exclusive services to their members, such as higher pay or better work conditions, because contracts negotiated by the

unions are extended also to nonunionized workers. This reduces the incentive to become a member of a trade union because one can free-ride on the services provided by the organizations. Unions are stronger in countries adopting the Ghent system (from the Belgian town where the first agreement was signed envisaging an active role of unions in the running of unemployment benefits) because membership qualifies a person for income support in case of job loss.

3.1.2 Coverage and Excess Coverage

A better measure of the bargaining power of trade unions is therefore provided by the coverage of trade unions, that is, the percentage of the eligible workforce—employees with bargaining rights—whose contract is regulated by the collective agreements signed by the unions. Unfortunately, there are no good cross-country-comparable measures of coverage. The first two columns of table 3.1 report data drawn from national surveys of employers or of membership in employers' associations involved in collective agreements. The third column displays membership rates. It is noticeable that in many countries coverage is much larger than union density.¹ For instance, in France, the country of the general strike, 9 workers out of 10 have their wage negotiated by trade unions, but less than 1 out of 10 is a member of a trade union.

This excess coverage of trade unions is a rather recent phenomenon. It is the result of (1) the decline in union density rates and (2) the presence in many countries of laws or practices that extend coverage beyond membership.

These regulations may themselves contribute to increasing the gap between coverage and membership because there are fewer incentives for workers to join a union, given that they benefit, in any event, from the wages negotiated by the unions (these extension clauses exacerbate the so-called free-rider problem of unions). Unions supporting the legal extension of the coverage of agreements are, in some sense, operating for their self-destruction.

In order to cope with this free-rider problem, unions in a number of countries have begun providing services exclusively to their members, for example, retirement counseling and tax advising (de facto outsourcing of government functions, as in Belgium and Italy), training and retraining (examples in Denmark, France, Germany, and Italy), or even financial services (the Netherlands). In countries that have adopted the Ghent system, unions are directly involved in the administration of unemployment benefits (Boeri, Brugiavini, and Calmfors, 2001).²

¹ In some countries, e.g., the United Kingdom, coverage is smaller than membership because collective agreements are applied only if a majority of workers in a firm are members of a union.

² Bockerman and Uusitalo (2006) document that membership in the Finnish trade unions declined by about 10 base points after the introduction of an unemployment benefit system competing with the system run by the unions and not imposing membership as a requirement for receiving income support in the case of job loss.

Table 3.1 Membership, Coverage, and Excess Coverage at the Turn of the Twentieth Century

	% Workers in firm joining employer association (%) (1)	% Workers covered by collective agreements (%) (2)	Workers joining trade unions market sector (%) (3)	Excess coverage (2) – (3)	Centralization	Coordination
Austria	96	97	34	63	3	4
Australia	–	80	35	45	2	2
Belgium	72	82	44	38	3	4
Canada	–	35	36	–1	1	1
Denmark	48	52	68	–16	2	4
Finland	58	67	65	2	5	5
France	74	75	10	65	2	2
Germany	72	80	25	55	3	4
Italy	40	81	36	45	2	4
Netherlands	79	79	19	60	3	4
Norway	54	62	44	18	4	4
Portugal	34	80	30	50	4	4
Spain	72	67	16	51	3	3
Sweden	56	72	77	– 5	3	3
Switzerland	37	50	22	28	2	4
United Kingdom	54	35	19	16	1	1
United States	–	13	10	3	1	1

Sources: Ebbinghaus and Visser (2000); Boeri, Brugiavini, and Calmfors (2001); OECD (2006a).

Notes:

Centralization:

1 Company and plant level predominant.

2 Combination of industry and company/plant level, with an important share of employees covered by company bargains.

3 Industry level predominant.

4 Predominantly industrial bargaining, but also recurrent central-level agreements.

5 Central-level agreements of overriding importance.

Coordination:

1 Fragmented company/plant bargaining, little or no coordination by upper-level associations.

2 Fragmented industry- and company-level bargaining, with little or no pattern setting.

3 Industry-level bargaining with irregular pattern setting and moderate coordination among major bargaining actors.

4 (a) Informal coordination of industrial and firm-level bargaining by (multiple) peak associations; (b) coordinated bargaining by peak confederations, including government-sponsored negotiations (tripartite agreements, social pacts), or government imposition of wage schedules; (c) regular pattern setting coupled with high union concentration and/or bargaining coordination by large firms; (d) government wage arbitration.

5 (a) Informal coordination of industry-level bargaining by an encompassing union confederation; (b) coordinated bargaining by peak confederations or government imposition of a wage schedule/freeze, with a peace obligation.

3.1.3 Coordination

Table 3.1 also provides information on the degree of centralization of collective bargaining, that is, the level at which contracts are negotiated, following a fivefold classification provided by the OECD (ranging from 1, which stands for firm level, to 5, which stands for national level). While in some countries (e.g., the United Kingdom or the United States) negotiation occurs mainly at the level of the firm, in continental and southern Europe it is more frequently carried out at the industry level, and there have been many “social pacts” implementing income policies (imposing ceilings on wage growth) at the national level. Even decentralized bargaining can be coordinated at the industry or national level by unions. Thus it is useful to complement information on the institutional level of bargaining with measures of the degree of coordination in collective bargaining, reported in the last column on the right-hand side of table 3.1. It is also important to distinguish between *formal centralization*, which has to do with the actual level at which wage contracts are signed and the existence of parallel union (or employer) organizations, and overall or *implicit coordination*, which could result also from informal (tacit) coordination between independent unions and employers.³

Sometimes hours of strike or the so-called wage share (the ratio of the total wage bill to GDP) are used as measures of unions’ power. However, a powerful union uses strikes only as a deterrent, and wage shares are outcome measures that do not necessarily correspond to the objectives of trade unions, maximizing wages only of their members rather than of the population at large. Nevertheless, both hours of strikes and wage shares in Europe display trends broadly similar to those of union membership. In particular, in the EU-15 the number of working days lost as a result of labor conflicts dropped from 85 million in 1979 (at the peak of union membership) to less than 7 million at the end of the 1990s (Boeri, Brugiavini, and Calmfors 2001), while wage shares declined by about 10 percentage points over the same time span. Importantly, wage shares did not decline in the 1980s and 1990s in the United States (Blanchard 1998; Blanchard and Wolfers 2000), which experienced a decline in union militancy similar to that in Europe.

3.2 Theory

Bargaining models offer a useful characterization of the effects of union activity on employment and wages. The implications of these models are dependent on the objectives of unions, which are themselves related to their membership when we go beyond the case of workers with identical preferences and productivity. Hence

³ For alternative measures of coordination, see Elmeskov, Martin, and Scarpetta (1998) and OECD (2004), as well as Nickell and Layard (1999).

we first discuss collective bargaining models, assuming identical union members, and then models of endogenous membership.

3.2.1 Collective Bargaining

The effects of unions on wages and employment can best be characterized as outcomes of a collective bargaining process involving, on the one hand, unions and, on the other hand, organizations of employers. The outcomes of this process depend on the bargaining power of the two parties and on the scope of bargaining, notably on whether it involves only wages or both wages and employment levels.

A standard (and realistic) characterization of collective bargaining is one in which unions and employers' organizations bargain over wages, which are then taken as given by individual employers who have the *right to manage* (Nickell and Andrews 1983) their firms, choosing the employment levels that maximize the profits of the firm, given the wage agreed at the collective bargaining table. It is assumed that the union is composed of identical workers, having the same reservation wage and being equally productive for their employers. This means that when labor is not unionized, labor supply is flat at the level of the reservation wage. The union maximizes the utility of these identical workers, while employers maximize profits. The outcome of the process depends on the bargaining power of unions. The latter is captured by a parameter (like β in box 3.1) ranging from 0, the competitive equilibrium case, to 1, where the union is instead a *monopoly union* (Dunlop 1944) that unilaterally sets wages.

Box 3.1 Bargaining over Wages (Right to Manage)

Consistent with bargaining theory (Binmore et al. 1986; Osborne and Rubinstein 1990), the outcome of a right-to-manage bargaining process is given by the maximization of the Nash (1950, 1953) product of the surplus of workers and employers over the no-agreement outcome (the fallback or status quo option) weighted by the bargaining power of unions. Assuming that union members are risk neutral, that is, $u(w) = w$, the surplus of each union member is given by the difference between the bargained wage and the utility while being nonemployed or reservation wage, that is, $(w - w^r)$. For the union as a whole, this surplus will have to be multiplied by the number of active members (unemployed members obtain no surplus), determined by labor demand, $L^d(w)$, given the right-to-manage structure of bargaining. When no agreement is reached, the firm does not produce, making zero profits. Hence the surplus of the employer is simply given by the profits, that is, the difference between revenues and costs, $\pi(w) \equiv y[L^d(w)] - wL^d(w)$.

(continued)

Box 3.1 (continued)

The right-to-manage agreement then obtains the wage level that maximizes the Nash product

$$[L^d(w) (w - w^r)]^\beta [y(w) - wL^d(w)]^{(1-\beta)},$$

where β is the bargaining power of unions. The wage level that solves this problem is implicitly given by the first-order condition of the Nash product:

$$\frac{w - w^r}{w} = \frac{\beta}{\beta \varepsilon_w^d + (1 - \beta) \varepsilon_w^\pi},$$

where $\varepsilon_w^d = \left| \frac{\partial L}{\partial w} \frac{w}{L} \right|$ and $\varepsilon_w^\pi = \left| \frac{\partial \pi}{\partial w} \frac{w}{\pi} \right|$ are, respectively, the wage elasticity of labor demand and the elasticity of profits with respect to wages. Thus the larger the bargaining power of unions, the larger the surplus that union members enjoy with respect to being unemployed. This union wage markup is decreasing with the absolute value of the two elasticities of labor demand and profits, with the latter being more important the lower the bargaining power of unions. When β tends to zero, the markup goes to zero, indicating that workers are paid their reservation wage, as in the competitive (and the pure monopsony) equilibrium. Notice further that as β tends to unity (the union has all bargaining power), the markup is simply given by the inverse of the elasticity of labor demand. The wage set by the unions in this case is the monopoly union wage (denoted by the superscript u)

$$\frac{w^u - w^r}{w^u} = \frac{1}{\varepsilon_w^d}.$$

The above clearly states that the more elastic the labor demand, for example, brought about by stronger competitive pressures, the lower the markup obtained by the union.

This model predicts that the stronger the bargaining power of unions, the higher the wedge or *markup* imposed by unions over the reservation wage (see box 3.1) and the lower the resulting employment level. These predictions can be graphically represented as in figure 3.2. The right to manage implies that employers decide upon L ; hence the collective bargaining outcome must be on the downward-sloping labor demand curve, L^d . There is also a restricted range of wage (and employment) levels that can possibly be achieved. Wages cannot be lower than w^r ; otherwise nobody would work because the reservation wage, the utility of nonemployment, would be higher than the wage. This lower bound of wage negotiations corresponds to the competitive equilibrium level, w^* , because we are on both the labor demand and the (nonunionized) labor supply schedules.⁴

⁴ Notice also that the pure monopsony equilibrium extracts all surplus from the workers, making them indifferent between working and not working.

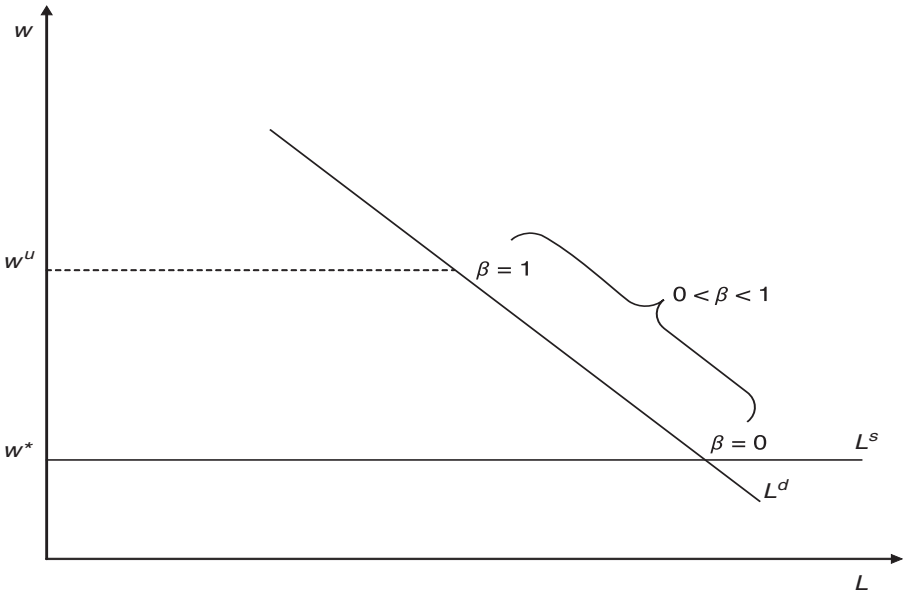


Figure 3.2 Right-to-Manage Outcomes and Bargaining Power of Unions

The upper bound of wage negotiations is the wage level, say w^u , that maximizes the surplus of union members, subject to being on the labor demand curve. The union has to take into account that higher wage claims, while increasing the surplus enjoyed by each worker over that worker's utility when unemployed (the reservation wage), reduce employment levels and hence the number of union members enjoying this surplus. As shown in box 3.1, the markup imposed by the unions over the competitive equilibrium wage is decreasing with the elasticity of labor demand and hence with the employment costs of wage hikes. In the intermediate cases (where $0 < \beta < 1$), the responsiveness of profits to wages, the costs imposed by unions on employers, will also matter in the bargaining outcome.

To summarize, under collective bargaining of the right-to-manage type, wages will be higher the stronger the bargaining power of unions, and the lower the responsiveness of labor demand and of profits to wages. These two elasticities are increasing with the degree of competition in product markets. Hence, consistent with the results discussed in chapter 1, stronger competition in product markets reduces the wedge between labor supply and demand introduced by labor market institutions. Technical annex 3.8 extends this result to the case where labor supply is not infinitely elastic (it is upward sloping) and where the case of zero bargaining power of workers coincides with the pure monopsony equilibrium characterized in chapter 2.

A right-to-manage structure of collective bargaining is inefficient in that employment is not involved in the bargaining process. Equilibriums must be on the

demand curve of firms, and this may rule out wage and employment outcomes that improve the situation of at least one bargaining agent without reducing the welfare of the other agent. In other words, there may be Pareto-superior equilibria that are disregarded in the right-to-manage model because of the assumption that only wages are bargained over. Graphically, the equilibria described in figure 3.2 may well not be on the *contract curve*, the curve denoting all Pareto-optimal allocations.

Efficient contracts (McDonald and Solow 1981; McCurdy and Pencavel 1986; Ashenfelter and Brown 1986), that is, outcomes of collective bargaining involving both wages and employment, will typically feature, for any given bargaining power of unions, more employment and lower wages than outcomes of bargaining over wages only.⁵ This is represented in figure 3.3, which displays again the right-to-manage outcomes (along the BC segment of the labor demand schedule), as well as the contract curve (the bold continuous line AB). As shown by the figure, the monopoly union outcome equates the slope of the labor demand to the slope of the indifference curve of unions, trading higher wages for less employment, but it is not tangent to any isoprofit curve, trading wages and employment levels from the standpoint of the firm. This means that the marginal revenues and costs of employment and wages are not equated among the two parties. This makes it possible to move, for instance, from C to A in figure 3.3, attaining a higher indifference curve of the union while being on the same isoprofit curve (generating the same profits as at C). A bargaining system is efficient only insofar as it equates the slope of the union indifference curves to the slope of the isoprofits. The loci along the AB segment are precisely defined by this condition.

Notice that in an efficient bargaining system, a higher bargaining power of union no longer implies less employment. In figure 3.3 the higher is β the closer is the outcome to A along the AB segment, thus higher levels of employment are attained than at any other outcome of right-to-manage bargaining. While the unions achieve higher levels of utility for their members by moving to the northeast of the figure, profits are increasing at lower wage levels. Hence the position along this contract curve depends on the relative bargaining strengths. Moreover, only the point B of the contract curve can be attained under a right-to-manage structure because B corresponds to the competitive equilibrium outcome, which is attained when unions have no bargaining power. Finally, the length of the contract curve (of the segment BA) depends on the extent of the rents that can be shared between employers and workers. The stronger the competition in product markets, the shorter the contract curve.

An interesting application of efficient bargaining models is when contracts are negotiated over wages and *standard* workweek hours (Booth and Ravallion

⁵ The only case where efficient bargaining involves less employment than the right-to-manage model is when workers are risk lovers, a case that is of limited empirical relevance.

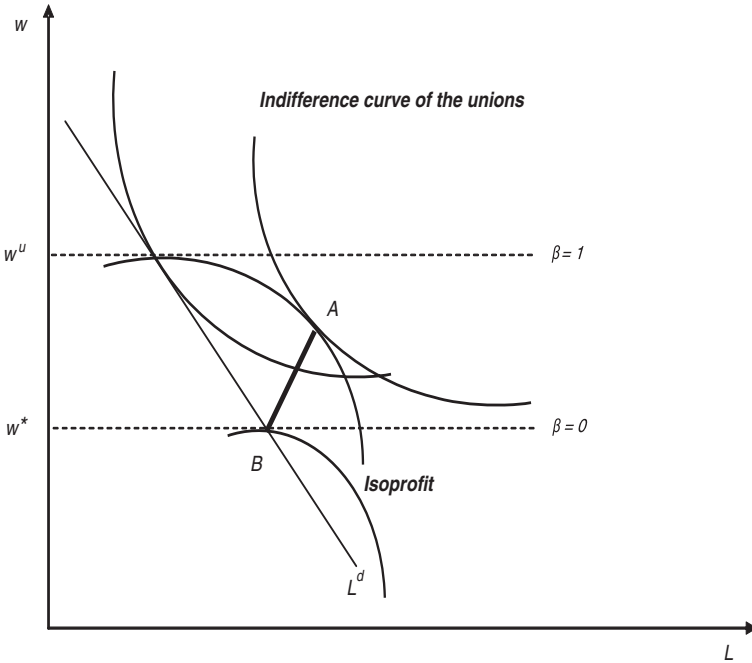


Figure 3.3 Efficient Bargaining Outcomes

1993; Contensou and Vranceanu 2000). Unions, in this context, can trade off higher *hourly wages* for shorter *working hours*. As in the models reviewed in chapter 5, the higher the value attributed to leisure relative to purchasing power by union members, the more likely that agreements will involve fewer hours worked. Reductions in working hours will not necessarily involve work sharing, that is, a “less hours-more workers” bargaining outcome. For work sharing to materialize, hourly labor productivity must be steeply decreasing with working hours, and unions should have a fairly low bargaining power so they can accept reductions in hourly wages or, at least, hourly wage increases that are lower than the increase in the value of the hourly labor product.

3.2.2 Endogenous Membership

The outcomes of collective bargaining previously discussed rest on the assumption that all union members have the same preferences and are equal also from the standpoint of firms. This makes it possible to characterize the objective of a trade union (TU) simply as the product of the surplus of each unionized worker and the total number of workers; that is,

$$TU = L(w - w^r).$$

When workers' preferences differ, maximization of the total surplus of union members involves interpersonal comparisons of utility. For instance, simply adding up the surplus of each individual worker implies giving the same weight to all union members, independent of their income level, or choosing more egalitarian objective functions of a trade union (e.g., a maximin rule, maximizing the well-being of the least well-off worker). By arbitrarily defining the function that maps heterogeneous preferences into a scalar measure of the union's objectives, one is therefore making a number of normative assumptions.

A positive approach to modeling the objectives of trade unions requires analyzing the internal decision structure of unions. Because trade unions are voluntary membership associations, their objectives ultimately depend on their membership, notably on the preferences of those paying the union dues. If the union has a democratic decision-making process, its objectives will be the outcome of a majority voting process. When members vote sincerely, there is a single decision (e.g., a wage platform), and preferences of individuals around this choice variable are single-peaked, then the objectives of the trade union will coincide with those of the median union member. Endogenous membership models are therefore useful also in characterizing the objectives of trade unions. They can be incorporated into monopoly union models by adding a membership equation that defines the identity of the marginal member who is indifferent between joining or not joining the union. It is then possible to locate the median union member and take that member's preferences as the objective function maximized by the monopoly union. This approach, however, implicitly assumes that union members are hired before nonunion members.

The models of endogenous membership also provide an explanation of why workers join a union when there are dues to be paid even when the results of union activity are available as a free good to all workers, independent of their membership in a trade union. This *union free-rider puzzle* is solved by pointing to the exclusive provision to union members of private goods such as on-the-job training (Acemoglu, Aghion, and Violante 2000), retirement and tax counseling (Booth and Chatterji, 1995), unemployment benefits, or access to soft-landing retirement plans (Brugiavini et al., 2001). An alternative explanation of union membership under regimes where a legal extension of collective agreements is provided relies on social norms (Akerlof 1980; Naylor and Cripps 1993; Corneo 1997b) that assign to union members a reputation, for example, for "good societal values."

The literature on membership also helps explain why unions typically pursue egalitarian wage policies. In addition to ideological factors, this egalitarianism can be motivated as the provision of insurance against wage fluctuations for risk-averse individuals (Agell and Lommerud 1992) or as a way to achieve savings on renegotiation costs and improve the relative position of low-skilled wage earners in the presence of strict employment protection legislation (Boeri and Burda 2008). Unions may also be involved in the enforcement of equal opportunity legislation,

particularly as women become more and more important in the membership of trade unions.

Not only does membership affect the objectives of unions, but also the wage platforms of unions affect membership. This is because, unlike minimum wages, unions act over the entire wage distribution, interfering with the way in which markets reward differences in productivity across workers. To give an example, in Italy in periods of double-digit inflation, unions sponsored a wage-indexing mechanism that granted the same absolute adjustment for inflation to all workers (the so-called “punto unico di contingenza”), deeply compressing the wage distribution. Suppose that inflation is 10 percent and that the uniform adjustment to inflation is 100 euros. Those initially earning less than 1,000 euros would increase or at least keep unaltered their real wage after the price change as a result of the automatic indexing mechanism. Those earning initially more than 1,000 euros would experience a decline of their real wage.

Often unions tend to pursue egalitarian wage policies that compress earnings differentials. This may crowd out the least skilled workers, located at the low end of the wage distribution, and reduce the skill premiums that would prevail in the absence of unions (see box 3.2 for a discussion of the sources of pay differentials in a competitive, nonunionized labor market and box 3.3 for theories of wage differentials in imperfect markets). This in turn may induce high-skilled workers to walk out of the union (in Italy, middle managers and high-skilled workers actually went on strike against the unions at one stage). Because the most skilled are not in the union and many of the least skilled are unemployed, membership becomes concentrated around intermediate-skill positions. A middle-skilled median union member is likely to vote for wage platforms increasing that member's pay relative to those workers with higher skills. To date, these interactions between union wages and membership have hardly been investigated by the literature.

Box 3.2 Compensating Wage Differentials

In the real world we observe large differences in pay across workers. The top 20 percent of wage earners typically receive five or six times the wage of workers located at the bottom 20 percent of the earnings distribution. The difference in pay varies considerably from country to country, suggesting that labor market institutions, such as unions, may play an important role in affecting the distribution of earnings that would prevail without these institutions. But why do markets price labor services differently?

In competitive labor markets wages may differ because of differences in the productivity of different workers. Labor is not a homogeneous factor of production. Because labor productivity combines characteristics of supply and demand, these differences will be associated with attributes of both the worker and the firm.

(continued)

Box 3.2 (continued)

To the extent that different workers have different preferences about jobs to be taken, a competitive equilibrium that maximizes the total surplus will involve a nondegenerate wage distribution, that is, some difference in hourly pay. Higher wages will compensate for greater effort or the difficulty of performing some tasks. Some workers will choose low-pay/low-effort combinations, while others will choose high-pay/high-effort combinations, and there will be no incentives to depart from this *compensating differentials* (Rosen 1974) allocation. Under these conditions, any institution that compresses wage differentials is bound to negatively affect employment and total welfare, although it may improve the relative position of some workers.

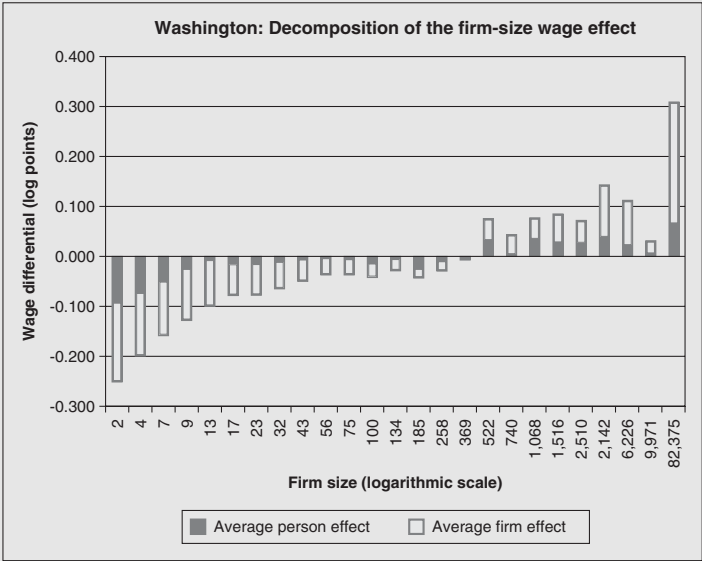


Figure 3.4 Decomposing Wage Differentials between Firms and Workers

Source: Abowd and Kramarz (2000).

The role played in wage differentials by workers’ heterogeneity (differences in the characteristics of individual workers) versus differences in firm characteristics can be characterized by drawing on the decomposition methods discussed in box 3.4. Matched employee-employer data, tracking characteristics of both workers and firms, allow us to disentangle workers versus firm heterogeneity as sources of wage differentials.

Figure 3.4 illustrates a similar decomposition for the state of Washington, performed by Abowd and Kramarz (2000). It documents that larger firms pay their workers more than small firms. This fact has been documented by many other studies and is called *firm-size firm-wage effect*. Importantly, this effect seems to be

(continued)

Box 3.2 (continued)

attributable mainly to firms versus workers' characteristics according to this study. Indeed, workers' heterogeneity (the black segment of the histogram) explains no more than 30 % of the total variation in pay, the remaining portion being attributable to firm effects.

Source: Abowd and Kramarz (2000).

Box 3.3 Wage Differentials in Imperfect Markets

In imperfect markets there are several alternative explanations for wage differentials. First, as discussed in chapter 2, when firms have *monopsony power*, employers maximizing profits offer an employment/wage combination that is lower than the competitive equilibrium. To the extent that in different industries there are different degrees of monopsony power, even when labor is equally productive across industries, market imperfections (monopsony power) may generate some wage dispersion. This explanation for pay differentials, however, is not entirely convincing. Indeed, if different industries priced the same type of labor services differently, then workers would move from low-paid to high-paid industries. This in turn would reduce the monopsony power of firms. In other words, there must be some frictions in the labor market that reduce the mobility of workers across jobs. These obstacles to mobility or barriers to entry prevent workers from moving across any pair of equal jobs offering different pay levels. Unions pursuing egalitarian wage policies may, in this context, play an efficiency-enhancing role if they complement markets in arbitraging away any difference in pay between workers-jobs matches yielding the same level of productivity. An important obstacle to labor mobility, and hence a convincing explanation for the presence of monopsony power, is the time lags associated with the matching of workers and vacancies. In the presence of such *matching frictions* there will be pools of jobseekers and of unfilled vacancies waiting to be matched to each other. If we suppose that all workers and vacancies are alike, an increase in the number of jobseekers will increase the probability that each vacancy is filled and reduce the probability that a jobseeker finds a job. Similarly, an increase in the number of vacancies will increase the job-finding rate while reducing the probability that a vacancy is filled. These *trading* (positive) and *congestion* (negative) *externalities* associated with the matching process make the institution-free equilibrium inefficient because it may create too many or too few vacancies and jobseekers at the equilibrium. A collective organization, like a union, may internalize these externalities, leading to a more efficient equilibrium.

Another mechanism that improves the efficiency of a decentralized equilibrium under matching frictions and allows wage differentials to arise at the equilibrium even when workers and jobs are perfectly homogeneous is *wage posting* (Moene 1997; Mortensen 2003). In this environment labor is perfectly mobile across recruitment pools, while employers announce the wage at the time of issuing the

(continued)

Box 3.3 (continued)

vacancy. Clearly, higher-paid vacancies will have a larger pool of applicants than relatively low-paid ones. Yet not all job applicants will apply for the high-paid vacancies because there are congestion externalities. As shown by Burdett and Mortensen (1998), some employers will offer relatively low wages that attract a relatively small workforce. Other employers will instead opt for a high-wage/high-employment combination. Employers offering higher wages are compensated by the longer duration of the jobs that they can offer. Workers in firms offering low wages will continue to search on-the-job for higher-paid jobs. Without on-the-job search there would be no dispersion in pay: wages would converge to the reservation wage of workers, as in the Diamond (1971) paradox. Overall, under wage posting the equilibrium involves a positive correlation between firm size and firm wage (like the one documented in box 3.2), as well as a positive wage/tenure profile. Both correlations are strongly supported by empirical evidence, although the latter (the positive relation between wage and tenure) can also be explained by human capital (chapter 8) and efficiency wages (see box 10.4).

Limited information is generally present on both sides of the market. Not only are workers not fully informed about job offers, but also employers may find it hard to obtain information about the number and characteristics of workers applying for jobs. This information problem may also give rise to wage differentials unrelated to workers or job heterogeneity. In particular, recruitment policies of firms may draw on signals that can be altered at will by workers. Educational attainments can be one of these signals, as discussed in chapter 8. Alternatively, employers in their recruitment policies may use information about the average quality of applicants in different pools (a town, an ethnic group, a group of workers with the same educational attainment) in selecting applicants. An optimal recruitment policy in this case may involve offering lower wages to workers coming from pools with lower average ability. This means that workers who are equally productive end up being paid differently simply because they come from different recruitment pools. This phenomenon is called *statistical discrimination*.

Sources: Burdett and Mortensen (1998); Moene (1997); Mortensen (2003); Rosen (1974).

3.3 Empirical Evidence

There are two main avenues of research in the empirical literature on the effects of unions on the labor market:

1. Estimates of the effects of unions on wages of members vis-à-vis non-members, that is, *union wage gaps*, and on the entire wage distribution, drawing mainly on individual data
2. Estimates of the effects of union density and bargaining coordination on employment, unemployment, and inflation, drawing mainly on macroeconomic time series

3.3.1 Effects of Unions on Wages

Economic theory suggests that unions increase wages above the reservation wages of individuals, extracting rents, if any are available, from employers. This prediction inspired an extensive empirical literature estimating the so-called union wage premiums, drawing on micro data.⁶ To bring the theory to the data, one has to acknowledge that all workers and jobs may not be equally productive and find appropriate controls for sources of heterogeneity in pay that are independent of unions' presence (e.g., the compensating wage differentials discussed in box 3.2).

The union wage gap is therefore estimated via cross-sectional regressions of wage equations of the type (where we omit error terms for simplicity)

$$\ln w_i = a + \beta_m D_i + X_i' \gamma,$$

where D_i is a dummy variable denoting membership in a trade union (it takes the value 1 when the individual is a member of a trade union and 0 otherwise), and X is a vector of personal characteristics affecting wages, such as age, educational attainment, and tenure on the job. If we denote by w^u and w^n mean wages of union and nonunion members, respectively, the estimated *union wage gap* is given by the parameter β_m ,

$$\frac{w^u - w^n}{w^n} \approx \ln w^u - \ln w^n = \beta_m.$$

Some authors also make D_i interact with personal characteristics so that unions are allowed to affect not only the average wage, but also the way in which age, gender, and seniority are rewarded by employers. The union wage gap indeed reflects both differences in the way in which personal characteristics are rewarded in the two segments and differences in the characteristics of union versus nonunion members (box 3.4 provides a simple methodology to disentangle these two components of the wage gap).

Box 3.4 Measuring and Decomposing the Effects of Unions on the Wage Distribution

The Oaxaca (1973) decomposition method has been widely used to decompose differences in *mean* earnings between unionized and nonunionized segments of the labor market into a component related to a pure *union wage premium effect* (differences in returns) and a component related to differences in the characteristics of union members versus nonmembers (differences in quantities). Suppose that

(continued)

⁶ See Booth (1995) for an excellent survey.

Box 3.4 (continued)

the mean wages of unionized and nonunionized workers can be written respectively as follows:

$$\bar{w}^u = \bar{X}^u \gamma^u$$

and

$$\bar{w}^n = \bar{X}^n \gamma^n,$$

where $\bar{X}^u$ denotes average personal demographic characteristics of unionized workers and γ^u the returns to these characteristics, while the superscript n denotes non members. Thus we can rewrite the union wage gap as the sum of the differences in quantities and differences in returns,

$$\Delta w = \bar{w}^u - \bar{w}^n = (\bar{X}^u - \bar{X}^n) \frac{1}{2}(\gamma^u + \gamma^n) + (\gamma^u - \gamma^n) \frac{1}{2}(\bar{X}^u + \bar{X}^n).$$

The Oaxaca decomposition implicitly assumes that changes in quantities do not affect prices and vice versa. Hence it does not take into account general equilibrium effects of changes in union wage premiums or in the number and characteristics of union members.

Another problem with the Oaxaca decomposition is that it cannot characterize effects of unions on the entire wage distribution. This is a problem because union activity is likely to exert important spillover effects on the entire wage distribution, for example, via substitution between labor and capital, as well as between skilled and unskilled labor. DiNardo, Fortin, and Lemieux (DFL henceforth) developed an ingenious method that makes it possible to estimate and analyze counterfactual earnings distributions rather than simply moments (like means) of these distributions. Their technique is a generalization of the Oaxaca decomposition. Essentially, their method involves three steps:

1. First, data from years with different levels of union membership, say t_0 and t_1 , are pooled together.
2. Next, the probability that each observation belongs to a given year, say t_0 , given individual attributes, is estimated via a probit model.
3. Finally, these estimates are used to reweight or map the distribution at t_1 into the distribution at t_0 so that observations that were relatively more likely in t_0 than in t_1 are weighted up and observations that are relatively less likely are weighted down.

Using this method, DFL conclude that the decline of unionization in the United States played an important role in explaining the collapse in the middle of the distribution of wages because unskilled workers slid back toward the lower tail of the wage distribution (see figure 3.5). Deunionization had instead a small effect on women's wages. Using similar methods, Card (2001) concluded that deunionization in the United States accounted for between 15 and 20 percent of the increase in wage inequalities, while it had no significant effect on gender pay differences.

(continued)

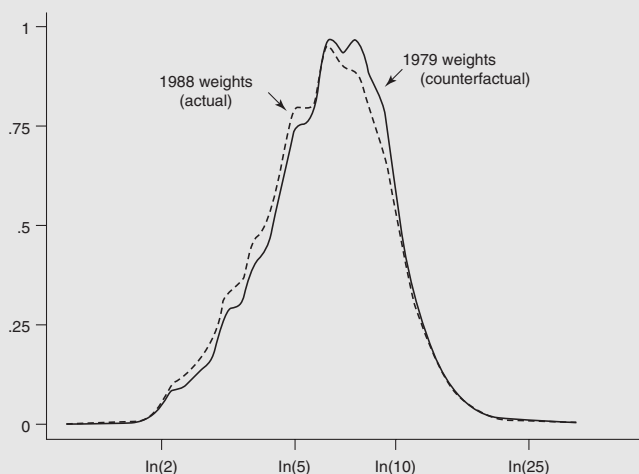
Box 3.4 (continued)

Figure 3.5 Estimated Effects of Deunionization on the U.S. Wage Distribution

The ingenious method developed by DFL is an extension of the Oaxaca decomposition, and its results should be interpreted with the same caution. In the words of DFL, “Calling the counterfactual density the ‘density that would have prevailed if individual attributes had remained at their 1979 level’ is a misuse of language. This distribution should rather be called the ‘density that would have prevailed if individual attributes had remained at their t_0 level and workers had been paid according to the wage schedule observed in t_1 ,’ since we ignore the impact of changes in the distribution of individual attributes on the structure of wages in general equilibrium.” The data used by DFL, as well as a program file allowing one to obtain the Oaxaca decomposition and the DFL method, are available at the following on our webpage (see link at <http://press.princeton.edu/titles/8771.html>).

Sources: Oaxaca (1973); DiNardo, Fortin, and Lemieux (1996).

As reported by Booth (1995), estimates of β_m range from 12 to 20 percent in the United States and from 3 to 19 percent in the United Kingdom. Estimates for other countries, which allow for legal extensions in the coverage of collective agreements, are not particularly informative because even workers who are not members are covered by union contracts. Although this literature suggests that union membership is associated with higher wages (the β_m coefficient is always

positive), estimates of the union wage gap can be very imprecise. Measurement errors can significantly affect the results because workers in the nonunionized segment may be less informed about their pay or some union members may not want to disclose their affiliation. A large union wage gap (a large β_m coefficient) estimated in these regressions may also reflect the fact that workers decide to be members of trade unions (technically speaking, *self-select* themselves into unions) in industries or firms where there are more rents to be split between employers and workers, and hence wages are higher. Self-selection of workers into unions (which are voluntary associations) generates a positive correlation between wages and membership that is due to reverse causality: it is because wages are higher in these jobs or occupations that workers are unionized rather than the other way around.⁷

In countries where the coverage of collective bargaining extends for beyond membership, it is difficult to assess the effects of unions on wages because there is no counterfactual wage distribution, no indication as to what wages would be without the unions. DiNardo, Fortin, and Lemieux (1996) developed an ingenious method to identify the effects of unions on the entire wage distribution and found that unions mainly affect wages of middle-skilled workers (see box 3.4). Blau and Kahn (1996), Kahn (1998, 2000), and Card (2001) also provided estimates of the effects of unionization (or deunionization) on the entire wage distribution. Even in countries where union membership and coverage of collective bargaining are closely associated, the presence of trade unions may exert spillovers on the entire wage distribution. This happens for a number of reasons. Unions may increase the bargaining position of nonunionized workers, alter the skill composition of the workforce and hence also the substitutability of labor with capital, or affect the sectoral composition of labor demand. If these spillover effects are important, the effects of union activity on wages can be evaluated only by considering the entire wage distribution. Similar problems arise also in the (much smaller) literature evaluating the effects of unions on productivity, hours of work, and employment. The empirical literature evaluating the effects of unions on the entire wage distribution typically finds that unions reduce wage dispersion, notably in countries with higher centralization/coordination of bargaining.

⁷ This endogeneity problem can be addressed, inter alia, by estimating, jointly with the wage equation, an endogenous membership equation, relating the decision to join a union to a number of covariates, including membership dues, industry-level labor productivity (or some measure of product market power to capture the size of the pie to be shared between workers and firms), and skill levels, in order to take into account that egalitarian wage policies of unions are bound to attract some skill groups more than others.

3.3.2 Bargaining Coordination, Union Density, and Unemployment

A number of studies have addressed the relationship between, on the one hand, union density and bargaining structure and, on the other hand, real wages and unemployment.⁸ This literature draws on the highly imperfect measures of bargaining structure discussed in section 3.1.3. These measures also exhibit a limited time-series variation. This is a serious problem because estimates draw on panel data seeking to explain variations both across countries and over time.

With these caveats in mind, a consistent finding in the macro empirical literature is that when broader coordination measures are used, a monotonic negative relationship between the degree of coordination and unemployment is observed, with higher coordination leading to lower unemployment. When centralization measures are instead used, a hump-shaped relationship is obtained (Calmfors and Driffil 1988), with low unemployment at low and high degrees of centralization and high unemployment with hybrid, intermediate bargaining regimes. Di Tella and McCulloch (2005), however, found more recently that greater bargaining centralization is associated with higher unemployment, a result that better helps understand the trend toward greater bargaining decentralization in the OECD area (see section 3.4.2).

It is interesting to compare the results on bargaining coordination with those using union density and coverage of collective agreements as right-hand-side variables. In general, these latter variables have lower explanatory power than bargaining coordination. The majority of the studies suggest that the difference between high and low density/coverage accounts for a smaller difference in unemployment than the difference between high and low coordination. A possible interpretation of these results is that changes in bargaining coordination are more important for macroeconomic outcomes than changes in union membership and coverage of union contracts. But it is not clear that one can treat density/coverage and bargaining coordination as independent of each other. It may be easier to achieve high coordination in the union sector when union density and coverage are higher because the benefits of coordination become larger when more employees are encompassed (Holden and Raaum 1991).

Overall, the (fragile) estimates from this macroeconomic literature suggest that the employment performance of an economy with both high bargaining coordination and high unionization is, *ceteris paribus*, superior to that of an economy with low coordination and unionization. At the same time, when coordination is lacking, better macroeconomic outcomes are observed under either centralized or decentralized regimes, with intermediate regimes offering the worst performance.

⁸ See OECD (2006a), chap. 3, for a survey of econometric evidence on the influence of coverage density and centralization coordination on equilibrium unemployment.

3.4 Policy Issues

3.4.1 Do Unions Increase Efficiency?

At least since the seminal works by Hirschman (1970) and Freeman and Medoff (1984), economists have usually characterized unions as twofold organizations, having both a good, efficiency-enhancing face and a bad, rent-seeking face. The *good face* of unions is associated with their function of collective *voice* of atomistic agents. Without such a voice, workers asking in vain for higher pay when productivity increases would have only the option of quitting the job and searching for another job with better pay (the *exit* option). Unions provide workers the option to continue to stay on the job and agitate for better pay, being at least as effective as when they exert the exit option, which is more costly because of the disruption of production associated with this mobility. By transmitting complaints, grievances, and demands, unions can also improve and correct the work relationship, making it possible to improve productivity in the firm. For instance, unions can force employers to provide more on-the-job training. Moreover, unions may help achieve higher efficiency by reducing transaction costs involved in individual bargaining.

In imperfect markets there are several second-best arguments for an efficiency-enhancing role of unions. For instance, unions can play the same role as a minimum wage in the presence of a monopsony (Robinson 1989). Indeed, historically, unions were created as a reaction to excessive monopsonistic power, and, as argued in chapter 2 and formally derived in technical annex 3.8, there can be efficiency gains in counteracting, at least up to some level, excessive bargaining power of employers. Unions may also provide protection through *ex post* redistribution against uninsurable labor market risk (Agell 2000; Jones and McKenna 1994), both by reducing wage differentials (hence earnings of workers in firms hit by adverse shocks) and by providing income in case of job loss, notably when unions are involved in the running of unemployment benefits, as in Ghent system countries. Clearly, it is not granted *a priori* that unions can be more efficient than the state in providing unemployment insurance. Finally, a number of circumstances in which unions, by compressing wage structures, improve economic efficiency are discussed in box 3.2.

The *bad face* of unions is their rent-seeking behavior. As suggested by figure 3.6, which draws on a representative survey of workers in four European countries to elicit membership in trade unions, unions are typically stronger in industries where there is less competition, like public utilities, and hence there are more rents to be split between workers and firms, while membership rates are nowadays particularly low in manufacturing, which is heavily exposed to international competition. According to the Eurobarometer Survey, in all countries of the EU unions are stronger in the public than in the private sector. In markets where competition erodes profits, unions cannot extract wages above the marginal

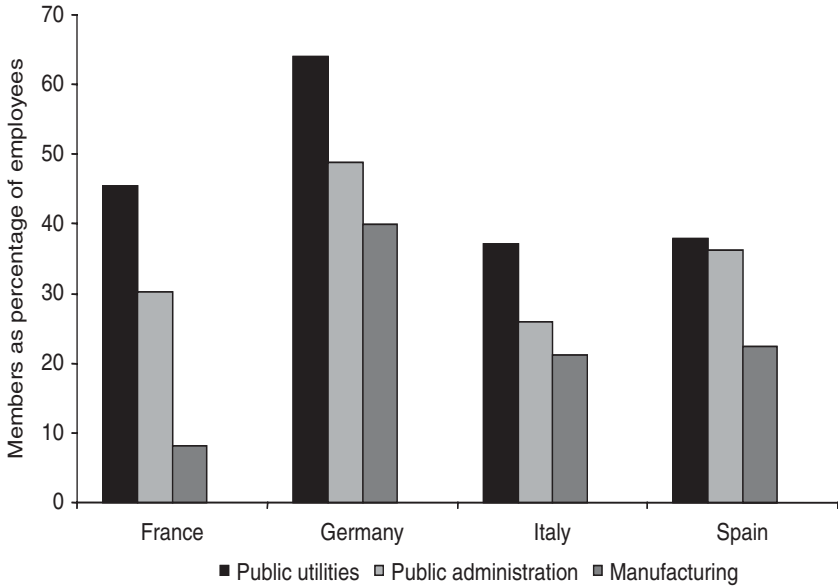


Figure 3.6 Union Membership in Different Sectors (2000)

Source: Fondazione Rodolfo De Benedetti survey for the 2000 Conference on Unions in the New Century.

product without forcing firms to exit or to shed labor. To put it another way, when competition is stronger, the segment AB in figure 3.3 becomes shorter. In Britain, one of the few countries with a long time series of microdata on union membership and firms' characteristics, the decline in union membership was stronger in industries and years where product market competition increased the most (Pencavel 2003). Blanchflower (2007) also reported that deunionization in Canada, the United Kingdom, and the United States was concentrated in the private sector, the union density rate among public-sector employees being five times larger than among private-sector workers.

The same mobility costs that confer monopsony power on firms (see chapter 2) may also confer bargaining power on workers and allow them to extract rents from their employers. This typically happens when employers have to make irreversible investments, for example, they need to train their workforce in order to make it productive. This irreversible investment creates a substantial asymmetry between the *incumbent* workers (those for whom the investment has already been carried out) and the *outsiders* who need to be trained before becoming at least as productive as the incumbents. This asymmetry can then be exploited by incumbents to extract some surplus from their employers. In other words, after the investment in training is made, workers can renegotiate their wage and obtain higher pay. This *holdup* problem (Williamson 1975; Grout 1984) can be particularly serious in innovative sectors requiring substantial investment in human capital at start-up. Unless unions

can credibly commit to a wage schedule negotiated *ex ante*, the threat of *ex post* wage renegotiation can discourage start-up of innovative firms and job creation.

Although redistribution in favor of low-skilled workers may be desirable from an equity standpoint, unions, by pursuing the interests of their members, often end up reducing employment opportunities for those who are not represented, that is, because they do not have a job. Egalitarian wage scales, in particular, may crowd out from the labor market the least skilled workers. Unions also generally support employment protection legislation (chapter 10) that makes entry and reentry to the labor market harder. This means that unions reduce the size of the labor market, leaving out new entrants and low-skilled workers.

Finally, the bad face of unions can materialize in their opposition to restructuring plans or to reforms aimed at increasing economic efficiency via reductions of rents in services less exposed to foreign competition. Coalitions of employers and workers' organizations often oppose regulatory reforms aimed at increasing competition in these sectors.

3.4.2 Should Collective Bargaining Be Decentralized?

These empirical results are consistent with the view that unions encompassing a very large set of workers can better take into account the macroeconomic effects of bargaining. In particular, centralized wage agreements can take into account the effects on inflation associated with excessive wage claims (Calmfors and Driffil 1988) and internalise aggregate demand externalities associated with consumer price inflation (Alesina and Perotti 1997; Soskice and Iversen 2000), as well as fiscal externalities (Flanagan 1999) related to the payment of unemployment benefits. However, bargaining systems in the OECD area have recently been evolving toward greater decentralization (OECD 2006c), and this trend is generally not imposed by governments, but chosen by unions and employers' associations (with the noteworthy exception of Australia, where the government actively encouraged decentralization). How can we reconcile this trend toward greater decentralization with the theoretical predictions and empirical findings previously reviewed on the advantages of centralized wage bargaining?

The presence of a hump-shaped relationship between bargaining centralization and unemployment suggests that countries with an intermediate degree of centralization (industry-level bargaining) can achieve a better employment (and inflation) performance by pursuing greater decentralization in wage setting. A strong theoretical argument in favor of decentralization can be made in the presence of large productivity differentials across industries and regions. Centralized wage agreements make wages unresponsive to these differences in productivity, increasing unemployment in the low-productivity industry or region. Another explanation for the superior macroeconomic outcomes of decentralized versus

hybrid bargaining structures is that the former can better accommodate incentives to achieve higher levels of efficiency, for example, via productivity-related pay.

Centralized or industry-level wage agreements can only replicate right-to-manage outcomes when unions agree with employers' associations on aggregate (or industry-level) wage levels, letting employment be determined by the individual employers, who are represented by their associations at the national or industry-level bargaining table. The right-to-manage model provides a more realistic description of how collective bargaining takes place under centralized or industry-level bargaining because it is, in any event, very difficult for employers' associations to commit their associates to the employment levels specified in the contract, and a choice of the efficient employment level would require local knowledge of the technology of firms and prescribe contingent contracts when the environment is uncertain. Thus efficient bargaining outcomes where both wages and employment are agreed on simultaneously can only be replicated at the plant level. Hence by decentralizing bargaining it is possible to achieve more efficient outcomes, potentially involving higher employment *and* wages, than in the right-to-manage model, as suggested by figure 3.3.

The decision on the bargaining regime does not truly belong to governments because it is agreed on by unions and employers associations. At most, governments can act as mediators in wage negotiations. Governments, however, may or may not allow the legal extension of collective agreements to nonunion members, and this indirectly affects the degree of centralization of bargaining. Some (rather weak) arguments in favor of this extension can be made on the ground that it can reduce wage negotiation costs and, potentially, labor disputes. Employers sometimes favor legal extension of egalitarian union-sponsored wage scales as a device to level the playing field and to reduce wage costs of skilled workers. However, extension enables voluntary associations representing the interests of a subset of workers to affect the entire workforce, while union leaders are accountable only to the members of the trade union. Hence legal extension of collective agreements can be deemed nondemocratic.

3.5 Why Do Unions Exist?

In spite of falling density rates, unions continue to exert significant influence over wage bargaining (and, more broadly, redistributive policies) in many countries, notably in Europe, because the coverage of collective bargaining did not decline together with membership. This widening excess coverage of unions suggests that simple measures of unions' strengths, which are typically based on membership or union presence in the workplace, can be quite misleading. In other words, unions' influence can be significantly larger than their presence.

Like any voluntary organization, unions exist because they are popular among some socioeconomic group. In this chapter we have shown who gains from the

presence of unions. These who gain are not necessarily low-skilled workers, who may actually be displaced by relatively high minimum wages imposed by the unions. Support of the egalitarian wage policies promoted by unions may come mainly from individuals with middle skill levels who can reduce their wage gap vis-à-vis highly skilled workers. Employers of skilled workers can also favor an extension of the coverage of union wages beyond the presence of unions at the workplace, notably when the presence of unions reduces negotiation costs, that is, when unions have not only a redistributive but also an efficiency-enhancing function. In any event, more theoretical work on the reasons for the large excess coverage rates observed in some EU countries is warranted.

Although trade unions are not keen to release information on the age of their members, household surveys suggest that the median age of a union member is increasing faster than that of the median worker. The fast aging of the median union member in some countries suggests that unions may be caught in a vicious circle of aging membership and reduced attractiveness among the young and active population. The share of retirees among union members is increasing everywhere; in Italy they are already a majority. This means that unions increasingly favor older people in intergenerational conflicts, for example, in the design of public pensions. Unless unions solve this intergenerational problem, they may be headed for the grave.

3.6 Suggestions for Further Reading

A thorough survey of the earlier theoretical and empirical literature on unions is provided by the book of Alison Booth on the economics of trade unions, which was published in 1995 by Cambridge University Press (Booth 1995). The first part of the Oxford University Press *The role of unions in the twenty-first century*, edited by Tito Boeri, Agar Brugiavini, and Lars Calmfors in 2001 (Boeri, Brugiavini, and Calmfors 2001) offers a more recent review of the literature on deunionization. An overview of the activities of unions around the world is offered by Richard Freeman's "What Do Unions Do?" in a symposium issue of the *Journal of Labor Research* published in 2005 and in Chapter 3 of the 2004 *OECD Employment outlook* (OECD 2004b).

3.7 Review Questions

1. What are the pros and cons of the various measures of the strength of labor unions provided by the literature?
2. What is the Ghent system?
3. Is there an efficient face of unions?
4. Why are unions stronger in industries where there is less competition in product markets?

5. What happens in a right-to-manage model when the bargaining power of workers increases?
6. Why is a right-to-manage bargaining system inefficient?
7. Why do unions pursue egalitarian wage policies?
8. How does competition affect efficient bargaining?
9. Why, in your view, does excess coverage exist?

3.8 Technical Annex: How Strong Should Unions Be in Order to Be Efficient?

Specify labor demand and supply as in technical annex 1.4. Collective bargaining in a right-to-manage environment involves the maximization of the product (rather than the sum, as in the perfect labor market case) of the surplus of employers and workers, that is, the Nash-bargaining rule

$$w = \arg \max \left(\left[\frac{AL^{1-\eta}}{1-\eta} - wL \right]^{1-\beta} \left[wL - \frac{1}{\epsilon+1} L^{\epsilon+1} \right]^\beta \right), \quad (3.1)$$

where the first term is the surplus of employers (profits) and the second the surplus of workers (the difference between the wage bill and reservation wages). Consistent with the framework developed in the technical annex to chapter 1, the fallback option of employers is zero (no production, hence no profits), and the fallback option of workers is the reservation wage represented by the constant-elasticity labor supply ($w = L^\epsilon$).

The two surpluses are weighted by the parameter β measuring the relative bargaining power of unions. Another interpretation of (3.1) is in term of a benevolent government caring about (functional) income distribution and assigning weights β and $(1 - \beta)$ to the welfare of workers and employers, respectively. More generally, problem (3.1) offers a qualitatively appropriate characterization of any situation where the contractual and institutional structure of the labor market addresses distributional concerns across two groups of agents (employers and employees), using nonmarket instruments (e.g., unions and employers' associations providing a voice to their members) to redistribute purchasing power within each group.

As discussed in this chapter, a realistic (albeit inefficient) bargaining regime is one in which employers and workers bargain only over wages' and then employers read from labor demand the employment level associated with this contracted wage (right-to-manage). Assuming then that employment is on the labor demand schedule, hence maximizing (3.1) with respect to w under the constraint that $L = (w/A)^{-\frac{1}{\eta}}$, we obtain

$$w^b = (\mu)^{\frac{\eta}{\epsilon+\eta}} (A)^{\frac{\epsilon}{\epsilon+\eta}} = (\mu)^{\frac{\eta}{\epsilon+\eta}} w^*, \quad (3.2)$$

where $\mu \equiv \left(\frac{1-\eta}{1+\varepsilon} + \beta \frac{\eta+\varepsilon}{1+\varepsilon}\right) \frac{1}{1-\eta}$ is the optimal markup imposed by collective bargaining over the opportunity cost of working, the superscript b denotes equilibriums with collective bargaining institutions, and w^* is the wage prevailing at the equilibrium without unions. This solution encompasses the case of a *monopoly union* (setting wages) faced by a right-to-manage (setting employment) employer: when $\beta = 1$, all weight is on worker welfare, and

$$w^b(\beta = 1) = w^u = \left(\frac{1}{1-\eta}\right)^{\frac{\eta}{\varepsilon+\eta}} (A)^{\frac{\varepsilon}{\varepsilon+\eta}} = \left(\frac{1}{1-\eta}\right)^{\frac{\eta}{\varepsilon+\eta}} w^*.$$

Notice that the monopoly union wage converges to the competitive equilibrium wage only when the aggregate labor demand is infinitely elastic (because η tends to 0). More generally, a monopoly union would impose on employers the wage

$$w^u = \left(\frac{1}{1-\eta}\right) w^*.$$

To put it another way, unions introduce a wedge between labor demand and supply that is decreasing with the wage elasticity of labor demand. In the cases where market power is on both sides (employers and workers) also the elasticity of labor supply matters. This helps explain why unions typically represent workers (e.g., prime-age males) with inelastic labor supply.

When instead all bargaining power is on the side of employers ($\beta = 0$), we have that

$$w^b(\beta = 0) = w^m = \left(\frac{1}{1+\varepsilon}\right)^{\frac{\eta}{\varepsilon+\eta}} (A)^{\frac{\varepsilon}{\varepsilon+\eta}} = \left(\frac{1}{1+\varepsilon}\right)^{\frac{\eta}{\varepsilon+\eta}} w^*;$$

that is, we go back to the pure monopsony case (hence the superscript m) characterized by (2.1) in technical annex 2.8. Like the monopoly union case, the pure monopsony equilibrium can replicate the competitive equilibrium wage only when labor demand is infinitely elastic (e.g., the firm has no monopoly power in product markets, and technologies are of the constant returns to scale type).

In all the more interesting cases where both labor demand and supply are inelastic, the competitive equilibrium outcome can only be replicated by an intermediate value of the bargaining-power parameter β . In particular, the equilibrium under Nash bargaining coincides with the competitive equilibrium when

$$\beta = \frac{\varepsilon}{\varepsilon + \eta} (1 - \eta), \quad 1 - \beta = \frac{\eta}{\varepsilon + \eta} (1 + \varepsilon),$$

because in this case $\mu = 1$ and the labor market generates the competitive equilibrium wage. In fact, the equilibrium in a perfect economy is supported by any

combination of weights such that the ratio of the profit share to the labor share is

$$\frac{1 - \beta}{\beta} = \frac{\eta}{1 - \eta} \frac{1 + \varepsilon}{\varepsilon}. \quad (3.3)$$

This condition is similar to the Hosios (1990) condition for efficiency when individual workers and jobs meet randomly according to a given matching technology, under constant returns.⁹ There is no reason to expect *a priori* that this condition is fulfilled when there is uncoordinated individual bargaining. In the latter case β can be simply interpreted as a subjective discount factor that reflects the relative impatience (hence weakness) of the two parties at the bargaining table. However, insofar as β is a reduced-form representation of an allocation mechanism different from perfect competition, it may be expected to react to changes in the relevant elasticities, ε and η . For instance, β can be interpreted as requiring that unions are not too strong when the labor demand elasticity is large in order not to reduce too much the size of the labor market and hence the pie to be shared between workers and firms.

To the extent that β is a (reduced-form) representation of politically supported equilibriums, this parameter may react to shocks (e.g., the competitive pressures arising from globalization that have been characterized in chapter 1) that increase the employment costs of redistributive institutions. For instance, collective bargaining institutions encompassing a rather broad range of interest groups may adjust wage claims to the new environment. Unions engaged in nationwide wage bargaining internalize the fact that unemployment will increase unless pay concessions are made. Small, decentralized unions may instead resist changes in their members' take-home pay. If every union follows the same policy, the outcome will be excessively high wages at the macroeconomic level, implying a bigger employment cost than with a nationwide union. This is consistent with the arguments originally developed by Calmfors and Driffill (1988) on the labor market effects of macroeconomic shocks under different bargaining structures. The adjustment of wage claims under collective agreements to environmental changes may take some time to materialize, depending on the frequency of contracts, but eventually will take place.

⁹ See technical annexes to chapters 10 and 11 for a simple matching model. It should be stressed that in that framework unemployment is present in equilibrium, but if the Hosios condition is satisfied, unemployment efficiently coordinates the search decisions of workers and firms in a frictional labor market.

This page intentionally left blank

4

Payroll Taxes

Historically, taxes emerged from seigneurial arrangements by which the king could obtain extraordinary revenues to meet unusual temporary conditions such as wars. From these extraordinary revenues medieval taxation developed, which ultimately became the fiscal basis for government expenditures (Ames and Rapp, 1977).

Payroll taxes consist of income taxes and social security contributions. Revenues from income taxation are an important source of public finance and are unrelated to the labor market. Social security contributions are directly related to the functioning of the labor market. Payroll taxes drive a *wedge* between the cost of labor to the firm and the net wage of the worker and therefore may have distortionary effects on the functioning of the labor market. Workers supply less labor than they would have with lower taxes, and employers create fewer jobs than they would have with lower taxes. However, taxes and social security contributions do not necessarily reduce the perceived income of the worker. Social security contributions may also be considered deferred consumption, for example, in the case of contributions to public pension programs. Individuals give up income today in exchange for a higher future income. Social security contributions are also a form of insurance against income fluctuations, for example, in the case of unemployment. The incidence of taxation is relevant too; the effects of taxes on behavior do not necessarily coincide with the point where the tax burden is put. It may be that taxes imposed on employers actually mostly affect workers' behavior. Finally, the effects of payroll taxes on the behavior of workers and firms cannot be studied in isolation because their effects are related to the use made of the tax proceeds. Therefore, this chapter also discusses social transfers to households and tax credits to workers and firms, and it discusses how payroll taxes and social transfers together influence work incentives.

There are clear overlaps between payroll tax institutions and other labor market institutions. The early retirement programs institution (chapter 7) also concerns elements of taxation, since high tax rates may induce workers to retire early. Furthermore, there is an overlap with unemployment benefits institutions (chapter

11) because the way unemployment benefits affect the labor market also depends on the structure of payroll taxes. Unions (chapter 3) are also important because, as will be discussed in more detail later, union bargaining affects the relationship among taxes, the structure of taxation, and labor demand and labor supply responses. Finally, there is an overlap between payroll taxes and active labor market policies (chapter 12) to the extent that incentives affect the behavior of unemployed workers in their search for jobs.

4.1 Measures and Cross-Country Comparisons

Employers are required to withhold payroll taxes from their employees' pay. These taxes are usually levied as a percentage called the *tax rate*. The *average tax rate* is the total tax paid divided by the total amount the tax is levied on; the *marginal tax rate* is the percentage tax paid on the next amount of money earned. In a *progressive tax system* the average tax rate increases with income; that is, the marginal tax rate is higher than the average tax rate. The *marginal effective tax rate* may differ from a marginal tax rate because the taxpayer may be in an income range in which that person is subject to a phaseout of some exclusion or deduction. Eligibility to several social programs is indeed conditional on passing an income and asset test (means-testing). As previously noted, payroll taxes consist of two components: taxes and social security contributions. Sometimes there is a distinction between security contributions paid by employees and social security contributions paid by employers, but this distinction is often considered irrelevant from an economic point of view. Tax systems are complex, and to summarize them by one particular number is impossible. Therefore, tax levels are often calculated for various types of situations.

Table 4.1 presents a cross-country overview of taxes and social security contributions for a single worker earning the average production worker (APW) wage.¹ Countries differ greatly in the overall tax burden, as well as in the composition of the two components. The average payroll tax ranges from a low of 17.2 percent in Korea to a high of 55.4 percent in Belgium. In most countries social security contributions are higher than income taxes; in Norway and the United States they are similar; and in Australia, Denmark, Iceland, and New Zealand income taxes dominate social security contributions. Table 4.1 also presents the marginal tax wedge, which also differs considerably from country to country. The lowest marginal tax rate is in Mexico (22.8 percent), the highest in Hungary (76.9 percent). In almost every country the marginal tax rates are substantially higher than the average tax rates. However, in countries like Sweden and Turkey the

¹ The average production worker wage is equal to the average annual gross wage earnings of adult, full-time manual workers in the manufacturing sector.

Table 4.1 Payroll Taxes and VAT Rates, 2005 (%)

	Average tax wedge		Total	Marginal tax wedge	VAT
	Income tax	SSC			
Australia	22.6	5.7	28.3	35.4	10
Austria	11.6	32.6	44.2	57.3	20
Belgium	21.4	34.0	55.4	66.4	21
Canada	14.9	16.7	31.6	40.7	7
Czech Republic	8.6	35.2	43.8	48.1	19
Denmark	30.2	11.0	41.3	49.2	25
Finland	20.1	24.5	44.6	54.9	22
France	10.8	39.3	50.1	55.8	19.6
Germany	17.3	34.6	51.9	65.1	16
Greece	4.4	34.4	38.8	54.1	18
Hungary	14.7	34.5	49.2	76.9	25
Iceland	23.4	5.6	28.9	39.7	24.5
Ireland	11.4	14.4	25.8	53.0	21
Italy	13.6	31.8	45.4	52.7	20
Japan	5.9	21.8	27.7	31.8	5
Korea	2.5	14.8	17.2	24.4	10
Luxembourg	11.1	24.1	35.2	52.4	15
Mexico	5.6	12.6	18.2	22.8	15
Netherlands	9.5	29.1	38.6	51.0	19
New Zealand	20.5	0.0	20.5	33.0	12.5
Norway	18.7	18.5	37.2	43.2	25
Poland	5.3	38.2	43.5	45.8	22
Portugal	8.2	28.1	36.3	47.1	19
Slovak Republic	6.9	31.4	38.3	44.4	19
Spain	10.7	28.3	39.1	45.5	16
Sweden	18.1	29.8	47.9	48.8	25
Switzerland	9.6	20.0	29.6	36.0	7.6
Turkey	12.7	30.0	42.7	44.5	18
United Kingdom	15.7	17.8	33.5	40.6	17.5
United States	14.6	14.5	29.0	34.0	–

Source: OECD tax database.

Notes: SSC is social security contributions; both average and marginal tax rates concern single persons without dependents who earn 100 percent of the APW wage; total average tax wedge is combined central and subcentral government income tax plus employee and employer social security contribution taxes as a percentage of labor costs defined as gross wage earnings plus employer social security contributions; tax wedge includes cash transfers. Note that in Australia, Canada, and New Zealand, VAT is known as GST (goods and service tax); the United States does not have VAT but sales taxes.

differences are small, indicating that the taxes are close to being proportional. Value-added taxes (VATs) are sometimes included in the calculation of the overall tax burden.² Table 4.1 shows that also for VATs there are big differences, ranging

² Nickell and Layard (1999), for example, advocate that the total tax wedge should also include consumption taxes (VATs) because it is the difference between product and consumption wages that matters for labor market behavior.

Table 4.2 Total Net Income and Marginal Effective Tax Rates of First and Second Earners (%)

	Total net income						Marginal effective tax rates		
	Employees			Unemployed					
	T1	T2	T3	T1	T2	T3	H1	H2	H3
Australia	76	86	86	25	46	56	69	63	66
Austria	71	89	89	39	64	70	99	21	41
Belgium	59	72	78	39	47	47	71	44	51
Canada	76	89	88	48	67	67	54	53	65
Czech Republic	76	86	90	38	53	65	99	31	37
Denmark	56	74	77	33	56	58	94	89	61
Finland	68	90	77	44	75	65	94	50	77
France	73	85	86	52	65	65	89	36	43
Germany	59	73	82	36	61	64	76	54	54
Greece	82	86	86	38	44	44	16	16	17
Hungary	71	92	90	31	50	49	38	18	37
Iceland	78	96	104	38	63	77	89	45	45
Ireland	84	98	101	43	58	74	88	34	42
Italy	73	84	89	38	50	54	-8	38	52
Japan	81	86	88	51	63	63	86	53	52
Korea	93	94	94	50	51	52	75	7	11
Luxembourg	78	102	102	66	92	92	84	62	62
Netherlands	67	81	77	48	63	61	88	42	41
New Zealand	79	80	80	47	50	53	77	56	62
Norway	71	86	82	47	70	60	87	27	36
Poland	68	69	71	30	38	52	87	65	68
Portugal	83	95	109	65	72	72	55	51	52
Slovak Republic	81	97	96	50	67	89	125	83	89
Spain	81	88	89	57	67	67	62	16	19
Sweden	70	92	79	56	83	66	100	34	47
Switzerland	80	91	92	57	75	89	99	14	22
United Kingdom	77	91	91	34	58	66	72	67	77
United States	76	85	89	43	47	51	46	52	52

Sources: OECD (2004a); OECD (2005b), box 3.2.

Notes: Family types: T1 = single person, T2 = lone parent with two children, T3 = one-earner couple with two children; included in total net income are earnings, social assistance, family benefits, housing benefits, income tax, own social security contributions, and in-work benefits; unemployed: previous earnings 100 percent of APW wage; initial phase of unemployment, previously full-time employed. Types of changes: H1 = one earner moving from inactivity to 67 percent of APW wage, H2 = first earner at 67 percent of APW wages, second earner moving from inactivity to 33 percent of APW wage, H3 = one earner moving from 67 percent to 100 percent of APW wage.

from a low of 5 percent in Japan and 7 percent in Canada to a high of 25 percent in Denmark, Hungary, Norway, and Sweden.

Table 4.2 presents information about the way in which the tax-benefit system affects the net income position of various types of employed and unemployed workers. The numbers represent net income as a percentage of the average wage

for an APW.³ As shown for a single worker, net income ranges from 93 percent of earnings in Korea to 56 percent in Denmark. Because of specific benefits or tax concessions to lone parents, household net incomes of working lone parents with average earnings in all countries are higher than net incomes of working singles, although in Greece, Korea, New Zealand, and Poland the differences are very small. Table 4.2 shows that for one-earner couples with two children net income can be even higher than gross income (Iceland, Ireland, Luxembourg, and Portugal). Finally, table 4.2 gives an overview of the way in which taxes and benefits affect the net income of unemployed workers in various family situations. As shown by the table, net incomes of unemployment benefit recipients vary considerably more across countries than those of employed workers. For a single worker, net incomes during the initial phase of unemployment range from 25 percent in Australia to 66 percent in Luxembourg.

For a single-earner couple with two children, net income in case of unemployment in the initial phase ranges from a low of 44 percent in Greece to a high of 92 percent in Luxembourg. As also shown in table 4.2, for a single earner moving from inactivity to 67 percent of the APW wage, the marginal tax rate in the Slovak Republic is more than 100 percent, while in Austria, the Czech Republic, Sweden, and Switzerland the marginal effective tax rate is about 100 percent. In these countries, for unemployed persons there are no incentives to search for a job if the expected wage is two-thirds of the APW wage. This contrasts sharply with Italy, where the marginal effective tax rate is below zero. Once one spouse is working, the other spouse is only marginally penalized when entering the labor market. Table 4.2 shows the effects of the marginal tax rate in case one spouse earns 67 percent of the APW wage and the partner is moving from inactivity to 33 percent of the APW wage. The highest marginal effective tax rate presented is 89 percent (in Denmark), while the lowest is 7 percent (Korea). Except for Italy and the United States, the marginal effective tax rates in the first situation are higher in every country than in the second situation. These characteristics of the tax-benefit system induce a polarization in both work-rich and work-poor families. If neither of the partners is working, the disincentives for one of them to accept a job are quite substantial, but if one of the partners is already working, the disincentives for the other partner also to find a job are much smaller. Finally, table 4.2 shows that the marginal effective tax rate for one earner moving from 67 percent to 100 percent of the APW wage ranges from a low of 11 percent in Korea to a high of 89 percent in the Czech Republic. It is clear that in many countries, in households where one person is already working, it is financially more interesting for the nonworking partner to enter employment than for the partner who is

³ The calculations include earnings, social assistance, family benefits, housing benefits, income tax, own social security contributions, and in-work benefits.

already working to work more. The only exceptions are Denmark, Japan, and the Netherlands.

4.2 Theory

Two issues concern payroll taxes as an institution. The first issue is whether the *presence* of taxes affects the functioning of the labor market. The answer is not necessarily. In a competitive labor market zero taxes are optimal. However, as discussed in section 1.4.3, in a noncompetitive market optimal taxes are generally nonzero. The second issue is whether the *structure* of the payroll taxes affects the functioning of the labor market. The answer is yes. It is not only the average tax rate that matters but also—and perhaps even more important—the marginal tax rate. Proportional tax rates generally have different effects than progressive tax rates. This second issue, the effect on the labor market of the structure of taxation, is the main topic of this theoretical section.

Payroll taxes affect both labor supply and labor demand. Labor supply is influenced along both the extensive and the intensive margins. First, it may be that the net wage is too low to encourage participation in the labor market, so the effective labor supply is reduced. The tax always decreases labor market participation because only the substitution effect matters. Second, it may be that the tax affects the number of hours supplied. Payroll taxes increase labor costs, and therefore labor demand will be reduced. Payroll taxes shift both the demand and the supply curves. Who pays the burden of payroll taxes depends on labor supply and labor demand elasticities (see also box 4.1).

The effects of the introduction of a payroll tax are depicted in figure 4.1. Initially, labor market equilibrium is at the wage/employment combination (L_0, w_0) . After the introduction of a tax, employment declines to L_1 , the wage costs for the firm are equal to w_d , and the net wages for the worker are equal to w_s . The total tax wedge is equal to $w_d - w_s$, of which the employer pays $w_d - w_0$ and the worker pays $w_0 - w_s$. The steeper the supply curve, the higher the share of the tax paid by the worker; the steeper the demand curve, the higher the share of the firm. As is spelled out in more detail in technical annex 4.8, the burden is shifted to the side of the labor market that is most inelastic. If labor supply does not respond much to changes in wages, employees pay a high proportion of the payroll tax. Because taxes increase the wage cost to employers and reduce the net wage of workers, taxes will reduce employment. The after-tax equilibrium will be inefficient because employment is not at the level that maximizes total gains from trade of labor services and introduces a deadweight loss, the triangle ABC .

Conditional on the total tax to be paid there is a choice between proportional or progressive taxes. In a progressive system the marginal tax rate depends on income, and the more hours an individual works, the higher the tax rate and thus the lower the marginal wage. As shown in table 4.1, in every OECD country the marginal

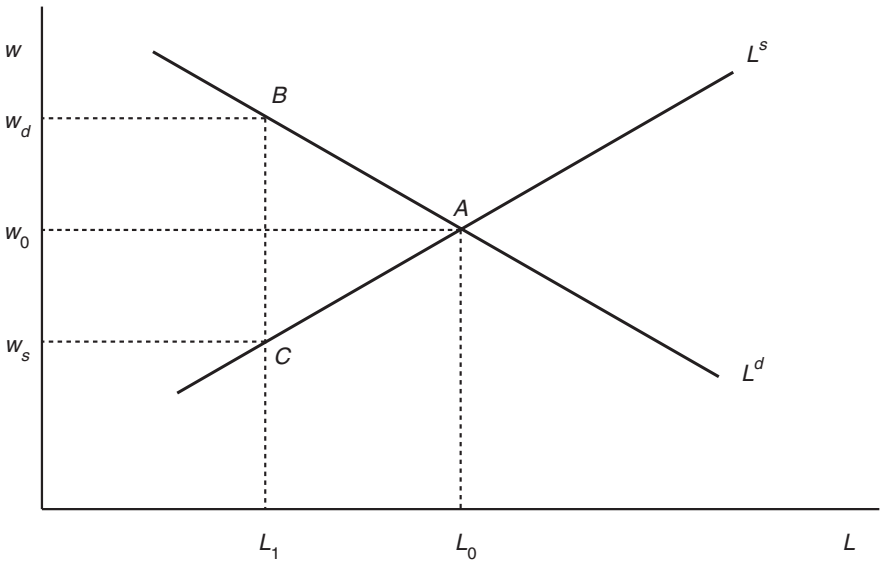


Figure 4.1 Payroll Taxes in Competitive Labor Markets

tax wedge is larger than the average tax wedge, indicating that all payroll taxes are progressive. Progressive tax systems are used for redistributive purposes. In many countries labor supply elasticity decreases with income. Since optimal taxation from the point of view of maximizing tax revenues requires that inelastic labor supply be taxed most, a progressive tax system is needed. Whether or not a progressive tax system induces an overall lower labor supply than a proportional tax system is unclear. This depends on the size of the substitution and income effects. If the substitution effect is larger than the income effect, a progressive tax system, with lower net marginal wages, will decrease labor supply. The opposite happens when the income effect dominates.

The impact of taxes on employment and wages depends not only on the nature of the labor market but also on the relationship between unemployment benefits and wages. In a competitive market where firms and workers are price-takers, labor demand is affected only by the tax level and not by the structure of the taxes. The effect of taxes on labor supply depends on the relationship between unemployment benefits and wages (see technical annex 4.8 for details). If unemployment benefits are indexed to prices, that is, real benefits are constant, the labor supply curve is upward sloping and an increase in taxes reduces labor supply. If unemployment benefits are indexed to wages, that is, the benefit replacement rate is constant, the labor supply curve has a steeper slope, and the effects of taxes on labor supply are mitigated; real wages are flexible and tend to absorb tax changes with little impact on employment. In some cases the labor supply curve is vertical; that is, tax rates do not affect labor supply.

The labor supply curve in a competitive market is determined by workers who take wages as given. Pissarides (1998) distinguishes three possibilities in addition to the competitive labor market: a labor market in which unions and firms determine the wage through bargaining (as in the models presented in chapter 3), a labor market where wages are determined by a bargain between firms and individual workers who search for a job, and an efficiency wage model in which the firm sets wages to avoid shirking of workers. If wages are determined through bargaining between firms and unions, the wage-setting equation determines labor supply. For unions to play a role, the wage-setting curve has to lie everywhere to the left of the competitive labor supply curve. Unemployment is an equilibrium phenomenon because it disciplines union wage demands. Taxes affect the wage equation because the unions and firms perceive that the tax paid depends on the wage choice. Because firms and unions take taxes into account, progressive taxation implies lower unemployment, while regressive taxation implies higher unemployment. Equilibrium employment depends on the structure of the employment tax system except in the case where taxes are proportional and unemployment benefits are fixed. With fixed benefits and proportional taxes, labor demand and the wage equation shift down by the same amount, so taxes are absorbed in wages without affecting employment. In a search model, because of heterogeneities in the labor market, it is costly for workers and firms to find each other, and once they have found each other, it is costly to break up the match. Because of these potential transaction costs, there are local monopoly rents associated with each job. Firms and individual workers bargain about this surplus, and because of the surplus there is unemployment. Also in this model, with fixed benefits, proportional tax changes are absorbed in wages, leaving employment unaffected. Finally, in an efficiency wage model the role of unemployment is to discipline workers and prevent shirking on the job. Given unemployment income, taxes do not influence the wage offer. Also, as in the competitive model, it is the amount of taxes that matters, not the structure of taxation. This makes the efficiency model closer to the competitive model than to the search or union-bargaining model.

The employment and wage effects of a tax cut depend on the slope of the wage-setting function. If this function is flat, a tax cut will lead to a large increase in employment. If the function is steep, tax cuts are absorbed by wages, and employment does not change much. These two situations are depicted in figures 4.2 and 4.3, respectively. In both figures the reduction in taxes causes the demand curve to shift from L_1^d to L_2^d . In figure 4.2 there is a large increase in employment and a small increase in wages. In figure 4.3 there is a small increase in employment and a large increase in wages. Workers use their market power to take most of the decrease in taxes.

Therefore, whether changes in the structure of taxation—that is, whether taxes are made more or less progressive—affect employment depends on the nature of the labor market. If wages are determined in a competitive market or in an

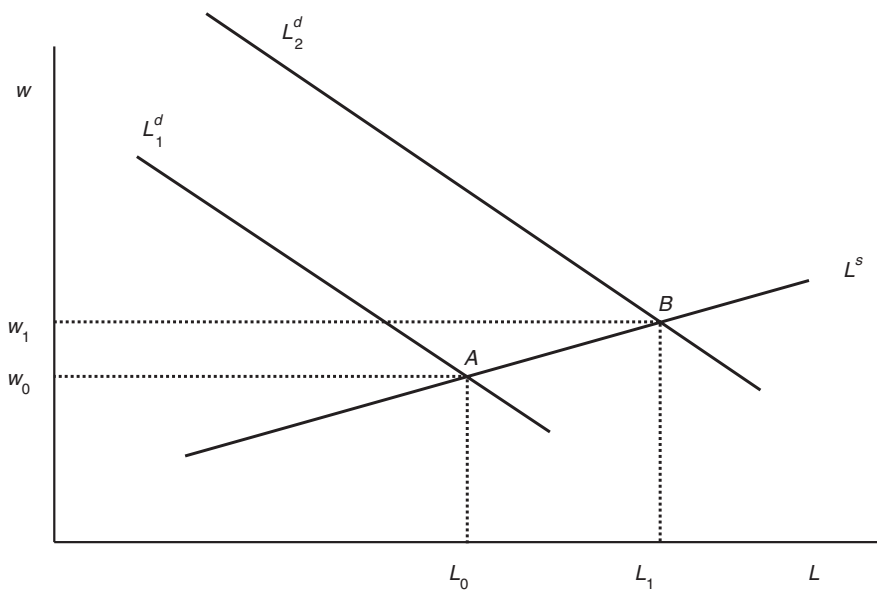


Figure 4.2 Tax Reduction in Competitive Labor Markets

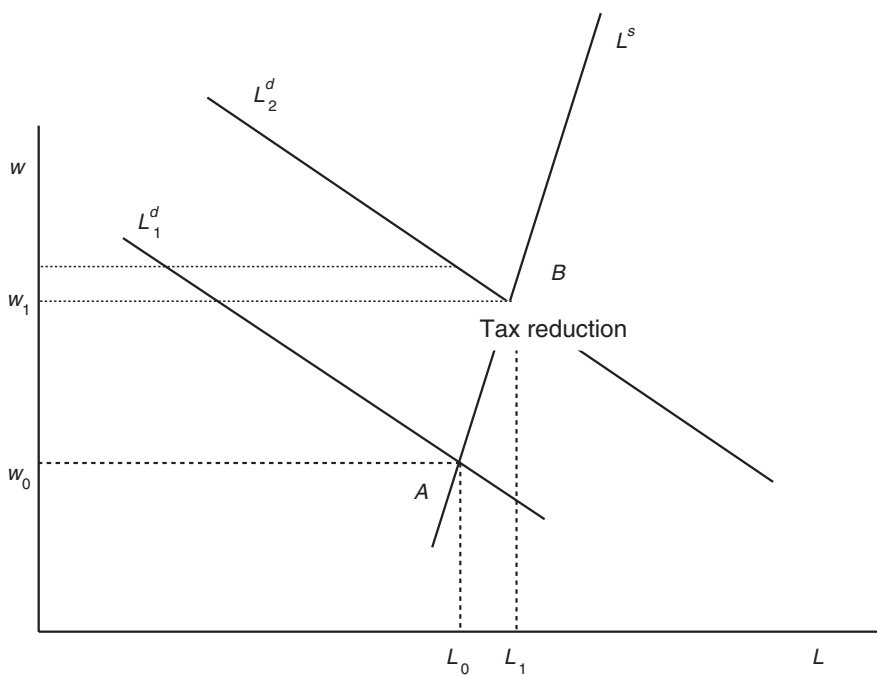


Figure 4.3 Tax Reduction in Noncompetitive Labor Markets

efficiency wage model setting, the structure of taxation is irrelevant. The only thing that matters is the level of taxation. However, in bargaining models, both individual search and collective union models, the structure of taxation is very important.

4.3 Empirical Evidence

A large empirical literature has studied labor supply and behavioral responses to taxes and transfers (see Blundell and MaCurdy 1999 for an overview). The labor supply of men is found not to be much affected by changes in tax rates, but the labor supply of women is usually found to increase if net wages increase. Eissa (1995), for example, uses the U.S. Tax Reform Act of 1986 as a natural experiment to identify the labor supply responsiveness of married women to changes in the tax rate. The tax reform reduced the top marginal tax rate by 44 percent (from 50 percent to 28 percent), but changed less the marginal tax rate for those further down the income distribution. Eissa analyzes the response of married women at or above the 99th percentile of the income distribution, using as a control group women from the 75th percentile of the income distribution. The tax effect is identified as the difference between the change in the labor supply of women with large tax rate reductions and the change in the labor supply of women with small tax rate reductions. Eissa finds evidence that the labor supply of high-income married women increased because of the tax reform, implying an elasticity with respect to the after-tax wage of approximately 0.8. At least half of this elasticity is due to a labor supply response along the extensive margin, that is, labor force participation. Several studies find relatively strong effects of tax reductions on participation of single parents (Eissa and Liebman 1996). Payroll taxes will only lower employment if taxes cannot be passed on to workers in the form of lower wages. If there is full compensation, employers are not confronted with higher labor costs, and higher taxes will not lead to lower employment. The question is whether taxes are fully incident on those who are formally liable to pay them. Gruber (1997) exploits a change in the financing of social insurance in Chile to assess the incidence of payroll taxation. The reduced costs of payroll taxation to firms appear to have been fully passed on to workers in the form of higher wages, with little effect on employment levels (see box 4.1 for details).

Box 4.1 The Incidence of Payroll Taxation

In a pure tax model workers do not benefit from taxes, but in reality payroll taxes may be used to finance programs that are beneficial to workers. If that is the case,

(continued)

Box 4.1 (continued)

a tax increase is buying the workers some benefits. In a static model payroll taxes affect labor supply (L^s) and labor demand (L^d) in the following way:

$$L^d = L^d(w(1 + t_f)), \quad (4.1)$$

$$L^s = L^s(w(1 - a_w t_e) + a_e w t_f), \quad (4.2)$$

where w is the pretax wage, t_f is the part of payroll taxes paid by the employer, t_e the part of payroll taxes paid by the worker, a_w the discounting of employee taxes by employees, and a_e the valuation of employer taxes by employees. Two extremes can be distinguished. First, in a pure tax model, $a_w = 1$ and $a_e = 0$; the worker does not benefit at all from taxes. Second, in a *full valuation* model, $a_w = 0$ and $a_e = 1$; all taxes are experienced as benefits to the worker.

The effect of a change in the part of employer taxes on the pretax wage can be expressed as

$$\frac{dw/w}{dt_f} = \frac{1/\epsilon - 1/\eta}{1/\eta - 1/\epsilon(1 - at_e)}, \quad (4.3)$$

where $1/\epsilon$ is the elasticity of labor supply and $1/\eta$ is the elasticity of labor demand. There are three situations in which a change in payroll taxes is fully shifted to wages ($\frac{dw/w}{dt_f} = -1$):

1. $\epsilon = \infty$; inelastic labor supply, that is, a vertical labor supply curve
2. $\eta = \infty$; perfectly elastic labor demand, that is, a horizontal labor demand curve
3. $a_w = 0$ and $a_e = 1$; the full valuation model; all taxes are considered benefits

In all three cases tax changes are fully compensated through a wage change, so a change in taxes should not have an effect on employment (L): $\frac{dL/L}{dt_f} = 0$.

Jonathan Gruber investigated the incidence of payroll taxation by exploiting a change in the financing of social insurance in Chile in 1981. Before 1981 most social insurance programs were financed by a substantial payroll tax. In May 1981 Chile shifted the financing of these social insurance programs from employer payroll taxes to general revenues. As a result there was a drop in the average payroll tax rate for manufacturing firms from about 29 percent in 1979/80 to 5.5 percent in 1984/85. Gruber found the following tax effects:

	$\frac{dw/w}{dt_f}$	$\frac{dL/L}{dt_f}$
Blue-collar workers	-0.9	0.2
White-collar workers	-1.4	-0.2

For both blue-collar workers and white-collar workers the wage effect is close to -1 , while the employment effects are close to 0. This implies that the reduced

(continued)

Box 4.1 *(continued)*

costs of payroll taxes have been fully passed on to wages, leaving employment unaffected. So there was full valuation, although it was not clear which of the three explanations for the phenomenon was most relevant.

Source: Gruber (1997).

In his overview of the labor market impact of taxes, Disney (2000) concludes that the labor supply of a regular full-time worker is probably inelastic to tax changes. He identifies four groups of workers where high tax rates may affect behavior: those in work with high incomes, those in work with low incomes and eligible for in-work benefits, those nearing retirement, and those considering entering the labor force. At the upper end of the wage distribution, high marginal tax rates may reduce labor supply and work effort. However, according to Disney, the effects are not sufficient to generate Laffer curve-type effects.⁴ At the low end of the income distribution, means-tested benefits discourage workers from participating in the labor market. For this reason, most OECD countries have introduced in-work benefits (discussed later). At the high end of the age distribution of the labor force, high tax rates have induced workers to retire early (see also chapter 6 on retirement schemes). At the low end of the labor force age distribution, entry to work may be discouraged by high effective marginal tax rates. This problem too has been tackled by creating in-work benefits.

4.4 Policy Issues

The structure of the tax-benefit system often does not generate strong incentives for low-income unemployed workers to find jobs. Policies aimed at “making work pay” introduce financial incentives to employment. These can consist either of subsidies to workers in the form of in-work benefits or of subsidies to employers in the form of tax credits or transfers to firms.

4.4.1 Negative Income Taxes or In-Work Benefits?

In a means-tested welfare system individuals without income receive a welfare benefit that is reduced one for one when the individual starts making money. In this system individuals are only tempted to enter the labor market if they can earn

⁴ The Laffer curve is promoted by the economist Arthur Laffer. Although he does not claim to have invented this concept, it was popularized with policymakers following an afternoon meeting with Dick Cheney, the 46th vice president of the United States, in which he reportedly sketched the curve on a napkin to illustrate his argument (www.wikipedia.com).

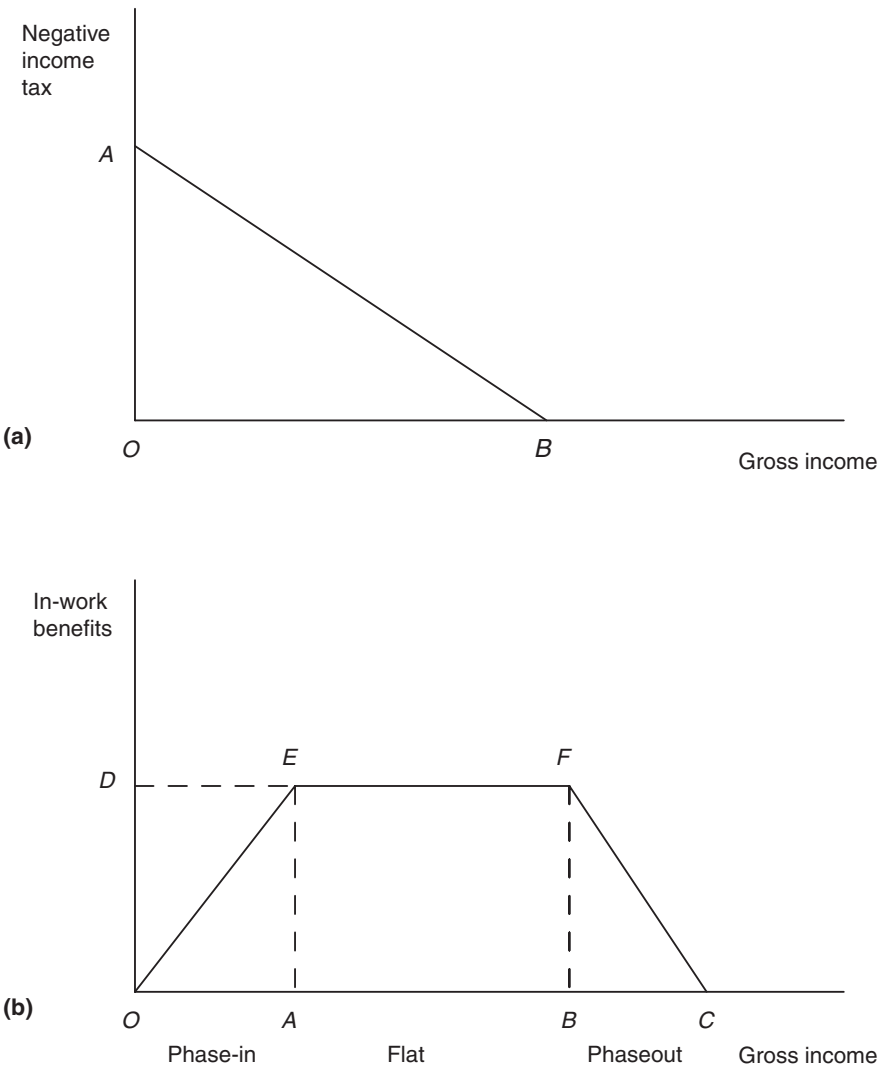


Figure 4.4 Design of the Negative Income Tax (a) and In-Work Benefits (b)

at least the amount of the welfare benefit. The negative income tax (NIT) was suggested by Milton Friedman (1962) as a proposal to reform the U.S. welfare system. With an NIT individuals get welfare benefits that are lower than the benefits under the original system, but benefits are withdrawn by a percentage of the increase in earned income in order to preserve financial incentives to work (see the upper part of figure 4.4). This provides incentives for some workers who want to work few hours, but other workers may be stimulated to reduce their working

hours.⁵ Despite several experiments, the NIT system has never been implemented in its original form.⁶

The design of an alternative benefit scheme, in-work benefits, is presented in the lower part of figure 4.4. Individuals receive benefits only if they work. The size of the benefits provided depends on the gross income of the worker. Over the gross income range *OA* the benefits initially increase to the maximum *OD*. This is the *phase-in* range. Then there is a flat region *AB* with constant in-work benefits. Finally, over the gross income range *BC* the in-work benefits are slowly reduced—the *phase-out* range.

Financial incentives to work are increased with in-work benefits. Card and Hyslop (2005) show that even for long-term welfare recipients, earnings subsidies have a positive effect on employment rates. However, temporary wage subsidies do not have a permanent effect (see box 4.2 for details).

Many countries have introduced some sort of in-work benefits with payments in the form of lump sums. Blundell (2001) offers a nice comparison of in-work in Canada, the UK, and the United States. The earliest scheme was the Earned Income Tax Credit (EITC) introduced in 1975 in the United States. The generosity of the modest program was increased in 1986, 1990, and 1993. The EITC is provided through the tax system rather than the welfare system. Eligibility for the EITC is awarded to all low-income families with children, irrespective of marital status. Receipt of EITC requires positive family earnings. Countries differ greatly in terms of benefit levels, withdrawal rates, time limits, hours limits, and targeting of programs. The entitlement criteria vary across countries. Whereas in Australia, Belgium, and Ireland entitlement is restricted to long-term unemployed individuals, in Canada, the Netherlands, and New Zealand all benefit recipients who find a job are entitled to an employment-conditional benefit. Except for Canada, hours of work requirements accompany the bonus payments: full-time employment in Australia and Ireland, part-time work in all other countries. Benefit levels vary widely across programs, with the United Kingdom's Working Family Tax Credit (up to 30 to 35 percent of average earnings for families with children and low income) and the Irish Family Income Supplement (paying 32.5 percent of the APW wage to working families with children) being the most generous. The U.S. system is less generous, with maximum payments of 13 percent of average earnings to families with two children. However, benefits are withdrawn faster in the United Kingdom (rate of 55 percent) than in the United States (rate of 20 percent).

⁵ NIT experiments in the United States in the late 1960s showed that labor supply responses of men along both margins of labor supply were small. For women, especially single female heads, and young workers, the behavioral response was higher and concentrated along the extensive margin. Participation elasticities were between 0.5 and 1. Studies of the EITC program, described later, show that the effect on participation of single female heads is substantial.

⁶ The main reason is that it was impossible to remove other transfers, which made the NIT system too expensive; see OECD (2005b), 149.

Box 4.2 In-Work Benefits in Action

In Canada in the early 1990s a new system of time-limited earnings subsidies for long-term welfare recipients was introduced. This system, the Self Sufficiency Project (SSP), was intended to help welfare recipients make a permanent break from program dependency. The subsidy scheme applied only to individuals who accepted full-time jobs and could last for up to three years, and once workers were offered the scheme, they had to begin to work within a year. The SSP was evaluated using a randomized design. Half of the long-term welfare recipients were offered the SSP, while the other half remained in the regular welfare system. Data were collected for six years in order to measure both the short-term and the long-term effects of the subsidy scheme. Card and Hyslop (2005) find that SSP had a large effect in the short run. The percentages of individuals who were on income assistance by years since random assignment were the following:

Years since program start	Control group	Program group	Effect of program
0.5	90.8	83.1	7.7
1	83.7	72.4	11.3
2	73.0	63.3	9.7
3	65.4	58.8	6.6
4	56.7	53.5	3.2
5	50.6	48.4	2.2
5.75	45.0	45.0	0.0

After the first 6 months of the experience there was already a difference of 7.7 percentage points between the treatment group and the control group. After 1 year the difference had increased to 11.3 percent. However, over time the effect faded away. After 3 years the difference between the two groups had fallen to 6.6 percent and after 5 years to 2.2 percent. After 69 months all subsidy payments ended, and the welfare participation rates of the two groups were equal. Once the financial incentives for some individuals to find a full-time job were gone, these individuals no longer behaved differently from individuals who did not get this financial stimulus. Card and Hyslop conclude that the SSP experiment offers little support for the idea that temporary wage subsidies can have a permanent effect on the labor market position of welfare benefit recipients.

Source: Card and Hyslop (2005).

The effects of in-work benefits depend on design features. Assessment of eligibility for in-work benefits can be done on the basis of family income or individual income. If eligibility is targeted on family income, the benefits go to families who need support the most, but there may be adverse effects on participation decisions

of secondary earners in couples. If it is targeted on individual incomes, this disincentive problem does not occur, but part of the benefits will go to high-income families.

In-work benefits may stimulate participation in the labor market for some workers, but the level of benefits provided still depends on income levels. Therefore, welfare systems reduce work incentives. In-work benefits are introduced to stimulate individuals without labor income to find a job by relating the provision of benefits to being employed. However, while, on the one hand, in-work benefits stimulate individuals to find a job, on the other hand, they reduce incentives to work longer hours (or earn higher wages). Therefore, whereas in-work benefits stimulate labor supply at the extensive margin, they reduce incentives at the intensive margin. The optimal income tax literature has focused almost exclusively on the intensive margin of labor supply (Saez 2002). The NIT provides the largest transfers to the lowest-income earners, who are presumably the most in need of support. NIT programs have adverse effects on labor supply along the extensive margin. In-work benefits programs do not provide income support to individuals with no earnings, but all earnings below a given threshold are partially matched by the government. In economic terms this is equivalent to negative marginal tax rates at the bottom of the income distribution. Relative to an NIT program, incentives to enter the labor force are enhanced with in-work benefits. This comes at the cost of the neediest people with no earnings, who receive no support. Furthermore, the in-work benefit has to be taxed away at some point further up the income distribution and thus imposes labor supply disincentives along the intensive margin for workers in the phaseout region. Saez (2002) concludes that if the behavioral response of individuals is along the intensive margin of labor supply, the NIT program is the optimal welfare program, with a large guaranteed income level that is taxed away at high rates. If the behavioral response is concentrated along the extensive margin of labor supply, then the optimal program is an in-work benefit with smaller guaranteed income level and transfers that increase with earnings at low income levels.

4.4.2 Tax Credits or Wage Subsidies?

Hiring decisions of firms may be influenced by payments to employers. These payments may be given as a direct transfer to the employer (wage subsidy) or as a tax credit. Wage subsidies may be provided as an incentive for firms to expand their employment (marginal wage subsidies). Tax credits are sometimes provided to firms to hire specific categories of workers. The trade-off between tax credits and wage subsidies is influenced by the costs of the payments. Tax credits do not increase public expenditures and are administered via the tax system, so there

is no need for a separate administration, as in the case of wage subsidies.⁷ The reduction in labor costs for specific types of workers may induce employers to hire more of these workers. In the United States the Targeted Jobs Tax Credit (TJTC) program began in 1979, targeting disadvantaged youth, the handicapped, and welfare recipients and providing employers with a tax credit lasting one year. The program was discontinued in 1995. There is evidence that this type of tax credits may have had stigmatizing effects on potential participants.⁸ Wage subsidies are sometimes targeted to low-wage employees and provided as employment subsidies to employers or as reductions to employers' social security contributions. Since the wage level is often the only qualifying condition, there could be an incentive for workers with high hourly wages to work part-time. An example of such a wage subsidy is the Dutch SPAK (*Specifieke Afdrachtskorting*), which was introduced in 1996 as a reduction of employers' contribution on low wages and phased out in 2003 over four years for cost-saving reasons. Evaluations suggested that the wage subsidy increased employment of low-skilled workers by 1 to 5 percent. Similar programs in France and Belgium also had significant employment effects (see Boeri 2005 for details). Tax credits and transfers to firms can generate deadweight losses if they subsidize existing jobs and job creation would have occurred anyway. The theoretical solution of subsidizing only marginal employment is difficult to use in practice.

4.5 Why Do Payroll Taxes Exist?

Payroll taxes exist to fund government expenditures. The social security contribution part of payroll taxes is directly related to the functioning of labor markets, for example, because it provides insurance to workers in case of unemployment. The payroll tax structure is important to ensure that workers have sufficient incentives to put effort into their job or search for a new job once they become unemployed. Although contributions to public pension programs are considered to be part of the tax wedge, it is not a priori clear whether these contributions are a tax or savings. Disney (2004) shows that the distribution over tax or savings of contributions to public pension programs is especially important for women. A higher tax component reduces employment rates of females, while a higher retirement savings component has the opposite effect. Employment rates of men are largely insensitive to tax rates. Tax rates are a significant factor in explaining differences in the amount of market work undertaken by the working-age population in different

⁷ Note that tax credits are not likely to be effective in case there is high tax noncompliance.

⁸ Burtless (1985) found that if tax credits were offered for disadvantaged welfare recipients, the probability that these workers were hired was in fact lower than if these tax credits were not offered.

countries. However, tax differentials explain only a minority of the market differentials, the majority being explained by other relevant labor market institutions (Nickell 2006).

4.6 Suggestions for Further Reading

Christopher Pissarides (1998) presents an analysis of the way taxes affect the functioning of the labor market under different assumptions concerning wage formation. Richard Blundell and Thomas McCurdy (1999) give an overview of the empirical literature on labor supply and behavioral responses to taxes and transfers. Chapter 3 of OECD (2005b) gives a detailed overview of the characteristics of in-work benefits systems in various countries. Tito Boeri (2005) discusses the components of an activating social security system in which in-work benefits and tax credits or transfers to firms are important.

4.7 Review Questions

1. How does a change in marginal taxes—with constant average taxes—affect labor demand and labor supply?
2. How do in-work benefits affect incentives of unemployed workers?
3. What happens during the phase-in and phaseout regions of in-work benefits?
4. How do supply and demand elasticities affect the distribution of the wedge between workers and firms?
5. How does the structure of payroll taxes affect the functioning of the labor market?
6. Why does the impact of taxes depend on the nature of the labor market?
7. What are the main purposes of negative income taxes and in-work benefits, and what are the essential differences between the two systems?
8. What is the trade-off between wage subsidies and tax credits?
9. Should contributions to public pension programs be part of the tax wedge?

4.8 Technical Annex: Taxes and Benefits in a Competitive Labor Market

A general structure for payroll taxes T is the linear tax,

$$T = tw + a, \quad (4.4)$$

where w is the wage the worker receives, t is the marginal tax rate on wages, and a is the per worker lump-sum tax or employment subsidy. The average tax rate $\frac{T}{w}$ is equal to $t + \frac{a}{w}$; the marginal tax rate $\frac{\partial T}{\partial w}$ is equal to t . Both a and t

are tax instruments where t is assumed to be positive or equal to zero, and a can be zero, positive, or negative. This linear tax subsumes various tax systems. Taxes are proportional if every worker pays the same tax rate irrespective of the level of the wage, in other words, if the average tax rate equals the marginal tax rate, which holds if $a = 0$. Taxes are progressive if the marginal tax rate is higher than the average tax rate, which holds if $a < 0$. Finally, the tax system is regressive if the marginal tax rate is lower than the average tax rate, which holds if $a > 0$.

Taxes may affect the demand for labor and the supply of labor. First, we consider the demand for labor. As in technical annex 1.4, in a competitive market in the short run (when capital is fixed), there is no loss in generality in assuming that the marginal value of a job is a decreasing (at a constant elasticity) function of the employment rate L , $AL^{-\eta}$, where A indexes labor productivity, and the index of the (inverse) labor demand elasticity, η , takes values between 0 (flat labor demand at A) and 1 (vertical labor demand at 1). Profits are maximized when the marginal value equals the opportunity cost of a marginal job, that is, the wage costs of the firm:

$$AL^{-\eta} = (1+t)w + a. \quad (4.5)$$

From this it follows that a tax increase—irrespective of whether it occurs through t or a —reduces labor demand. However, in a competitive market where firms are price-takers, labor demand is only affected by the tax level and not by the structure of the taxes.

The effect of taxes on labor supply depends on whether unemployment benefits are indexed to wages or to prices. Following Pissarides (1998), we assume that in a competitive labor market unemployment is time off work during which the worker receives unemployment benefits. The potential labor supply is normalized to unity. The actual labor supply depends on the preferences of the worker, which are assumed to be represented by a Cobb-Douglas-type utility function

$$V = c^\alpha (1-h)^{1-\alpha}, \quad (4.6)$$

where V represents utility, c represents consumption, h is working time, that is, labor supply, and α is a positive parameter with $0 < \alpha < 1$.⁹ If we assume that savings and wealth do not influence labor supply, consumption can be specified as $c = hw + (1-h)b$, where b , with $b < w$, represents the level of unemployment benefits. Thus the utility derived from leisure and labor supply becomes

$$V = (hw + (1-h)b)^\alpha (1-h)^{1-\alpha}. \quad (4.7)$$

⁹ Note that Pissarides (1998) uses a more general specification of the utility function.

In a competitive labor market the worker maximizes utility through labor supply ($\frac{\partial U}{\partial h}$), which gives the following first-order maximization condition:

$$\alpha(hw + (1-h)b)^{\alpha-1}(w-b)(1-h)^{1-\alpha} - (hw + (1-h)b)^{\alpha}(1-\alpha)(1-h)^{-\alpha} = 0, \quad (4.8)$$

from which, after some rewriting, we find the labor supply equation

$$h = \frac{\alpha w - b}{w - b}. \quad (4.9)$$

Concerning unemployment benefits, we can distinguish two situations. First, benefits are indexed to prices; that is, real benefits are constant. In this case, as shown by equation 4.9, the labor supply curve is upward sloping. Second, benefits are indexed to wages; that is, the replacement rate is constant. Rewriting equation 4.9 leads to

$$h = \frac{\alpha - b^w}{1 - b^w}, \quad (4.10)$$

where $b^w = \frac{b}{w}$ is the replacement rate. If the replacement rate is fixed, labor supply does not depend on the wage; that is, there is a vertical labor supply curve. All changes in the tax rate are absorbed in the wage, leaving unemployment unaffected. This is an extreme case and is related to the Cobb-Douglas specification of the utility function. With a more general specification also, when unemployment benefits are indexed to wages, the labor supply curve will be upward sloping, but the slope will be steeper than in the case in which unemployment benefits are indexed to prices.

Regulation of Working Hours

Working hours have been a topic of heated debate between unions and employers for a long time. May 1, 1886, was a day of strikes in the United States for the introduction of an eight-hour working day. To commemorate the working-hours struggle between unions and employers, May 1 was declared Labor Day, which is still a holiday in many countries. The discussion on the optimal length of the working day concerns wages, leisure, and productivity. In past decades there have been a number of clear trends in hours of work. Working hours per week have been declining slowly but have remained constant in many countries more recently. Average working hours go down not only through a reduction of standard working hours but also through part-time work. Part-time work has been increasing in most countries, especially for females.

The regulation of working hours institution involves two main issues. The first issue concerns regulation of standard weekly working hours; the second concerns the barriers workers face to choosing part-time work instead of a full-time job. Both issues concern decisions by workers and employers within the context of rules and regulations that originated from government interference or through collective agreements between unions and employers.

The regulation of working hours institution has clear relationships with collective bargaining and unions (chapter 3) because in many countries unions bargain about working hours. Furthermore, there is a relationship with family policies (chapter 7) because the presence of part-time work allows the combination of work and care. Finally, there is a clear relationship with employment protection legislation (chapter 10) because strong employment protection induces firms to adjust employment through the intensive margin rather than through the extensive margin. Furthermore, in many countries part-time jobs are also temporary jobs, and this makes adjustment of employment easier. Finally, adjustment of the workforce often involves fixed costs, which mainly consist of costs related to employment protection legislation, and this affects hours of work.

5.1 Measures and Cross-Country Comparisons

The usual measure of the intensive margin of labor supply is the average number of working hours per week. This can be averaged over all workers or can

be specified separately for full-time and part-time workers. Often the number of weekly working hours of part-time workers is roughly half the number for full-time workers. The choice concerning weekly hours of work is determined by workers' preferences, but the menu to choose from is determined by legislation and collective agreements. As will be discussed in more detail later, annual working hours are also strongly affected by the number of working weeks per year since in some countries workers have many more holidays than in other countries.

Table 5.1 presents an overview of legal restrictions on weekly working hours. In many countries the legal maximum of a *normal* working week is 40 hours; in Denmark the maximum is as low as 37 hours. There is wide variation in the maximum number of weekly *overtime* hours, from 2 in Spain to 16 in Switzerland. There is also wide variation in the maximum number of total (normal + overtime) weekly hours. In some countries this legal maximum is not the sum of legal normal hours and legal overtime hours because for a short period of time the maximum weekly overtime hours may sometimes be higher.¹ During overtime hours employers have to pay premiums that are usually in the range of 25 to 50 percent but in some countries may go up to 100 percent. Because of the overtime premiums, employers have an incentive to stick to normal working hours, but overtime hours allow employers to deal with temporary fluctuations in economic activity. The fourth column of table 5.1 shows that the normal weekly hours set by collective agreements are often substantially lower than the legal maxima of normal weekly hours.

A rough measure of the intensive margin of labor supply is the share of part-time work in total employment. Cross-country differences in part-time work are determined by the characteristics of part-time jobs, for example, wages and prospects for promotion or on-the-job training. In some countries unions oppose part-time jobs, which they consider to be inferior to full-time jobs. In other countries wages and social security arrangements for part-time jobs are similar to those for full-time jobs.

5.2 Theory

According to neoclassical labor supply theory, individuals choose the number of hours they want to work on the basis of the hourly wage rate and their preferences for leisure and income. The preferred number of hours may relate to working hours per day, working days per week, workweeks per year, and working years over the lifetime. Individuals supply labor to obtain an optimal combination of

¹ For example, in Austria the usual legal maximum of weekly overtime hours is 5, but during 12 weeks per year it is 10.

Table 5.1 Legislative Limits on Normal Weekly Hours of Work and Overtime Work

	Legal maxima on working hours			Bargained normal hours	Wage premium on overtime hours (%)
	Normal hours	Overtime hours	Maximum hours		
Australia	38–40	none	none	35–40	50
Austria	40	5	50	36–40	50
Belgium	40	10	50	38	50
Canada	40–48	none	none	35–40	50
Czech Republic	40.5	8	51	–	25
Denmark	37	none	48	37	50
Finland	40	5	45	37.5–40	50
France	39	9	48	39	25
Germany	48	12	60	35–39	25
Greece	40	8	48	40	25
Hungary	40	12	52	–	50
Ireland	48	12	60	38–40	25
Italy	48	12	60	36–40	10
Japan	40	none	none	40–44	25
Korea	44	12	56	–	50
Luxembourg	40	8	48	40	25–50
Mexico	48	9	57	–	100
Netherlands	45	15	60	36–40	–
New Zealand	40	none	none	40	–
Norway	40	10	50	37.5	40
Portugal	40	12	54	35–44	50
Spain	40	2	47	38–40	–
Sweden	40	12	52	40	–
Switzerland	45 or 50	16	61 or 66	40–42	25
Turkey	45	–	–	–	50
United Kingdom	none	none	none	34–40	–
United States	40	none	none	35–40	50

Source: OECD (1998).

Notes: Bargained normal hours refer to normal weekly hours set by collective agreements; legislative wage premiums for overtime hours concern the first overtime hours worked during the week. Luxembourg has a 25 percent rate for blue-collar workers and a 50 percent rate for white-collar workers; the Netherlands, New Zealand, Spain, and Sweden have no legislation on overtime hours wage premiums; the United Kingdom has a collectively bargained wage premium.

income and leisure; that is, they supply working hours until the marginal value of leisure equals the wage rate. In neoclassical labor demand theory employers are interested in hiring workers to put them to work for a number of hours at a given wage. Employers hire labor until the marginal value of an additional hour of work equals the wage costs. In reality the choice of working hours is often restricted to a limited set, most commonly full-time work, part-time work, and not working.

5.2.1 Regulating Weekly Working Hours

Shorter working hours may be introduced through mandatory laws or may be the result of bargaining between unions and firms. A reduction of the workweek is motivated for several reasons. Shorter working hours may increase the standard of living in line with preferences of workers. Sometimes work sharing is motivated as a tool to reduce unemployment: by reducing working time per worker, the number of jobs is supposed to increase. This, of course, is only true if the reduction of the working hours does not affect labor demand, that is, if there is a *lump of labor* that can simply be redistributed across workers.

The upper part of figure 5.1 shows the essence of this line of reasoning. We start in a competitive labor market, assuming that a firm requires a number of working hours as input to the production process. Initially there are no fixed costs of hiring a worker, and productivity does not depend on the number of hours per worker. In this simple setup the combination of workers and hours per worker does not matter; only the product of the two matters. The employer can move from one combination of hours per worker and number of workers to another without adjustment costs. The hyperbolic curve in the upper part of figure 5.1 shows the isohours curve, that is, all combinations of workers (L) and hours per worker (h) that generate the same number of working hours. If all workers work the same number of hours, h_1 , the firm needs L_1 workers. If in this setup the number of working hours per worker is reduced to h_2 , the firm will need L_2 workers: the working-time reduction has created new jobs for workers.

However, as Calmfors and Hoel (1989) argue, even if output is fixed, the reduction of working time may actually reduce labor demand and cause employment to fall. They indicate five channels through which this might occur. First, hourly wages may increase because workers want to protect their real income. Second, even if hourly wages do not change, wage costs per unit of time may increase because of the existence of fixed costs per worker. Third, labor productivity may fall if nonproductive time—start-up and finishing work—remains constant. Fourth, the costs of hiring new workers may rise relative to the costs of increasing working time when standard hours are cut. Finally, capital utilization may decrease if operating time is reduced parallel to working time.

It may be that the change from one combination of hours per worker and number of workers to another combination has a negative effect on the total number of hours; that is, the reduction of hours per worker may cause an inward shift in the isohours curve. This situation is depicted in the lower part of figure 5.1. Again the number of hours per worker is reduced from h_1 to h_2 , but because the total number of hours worked falls from $h_1 L_1$ to $h_2 L_2$, employment falls from L_1 to L_2 .

So far, we have ignored the fact that overtime hours often pay a higher wage. If every hour is paid the same wage, the iso-wage-cost curve for hours per worker

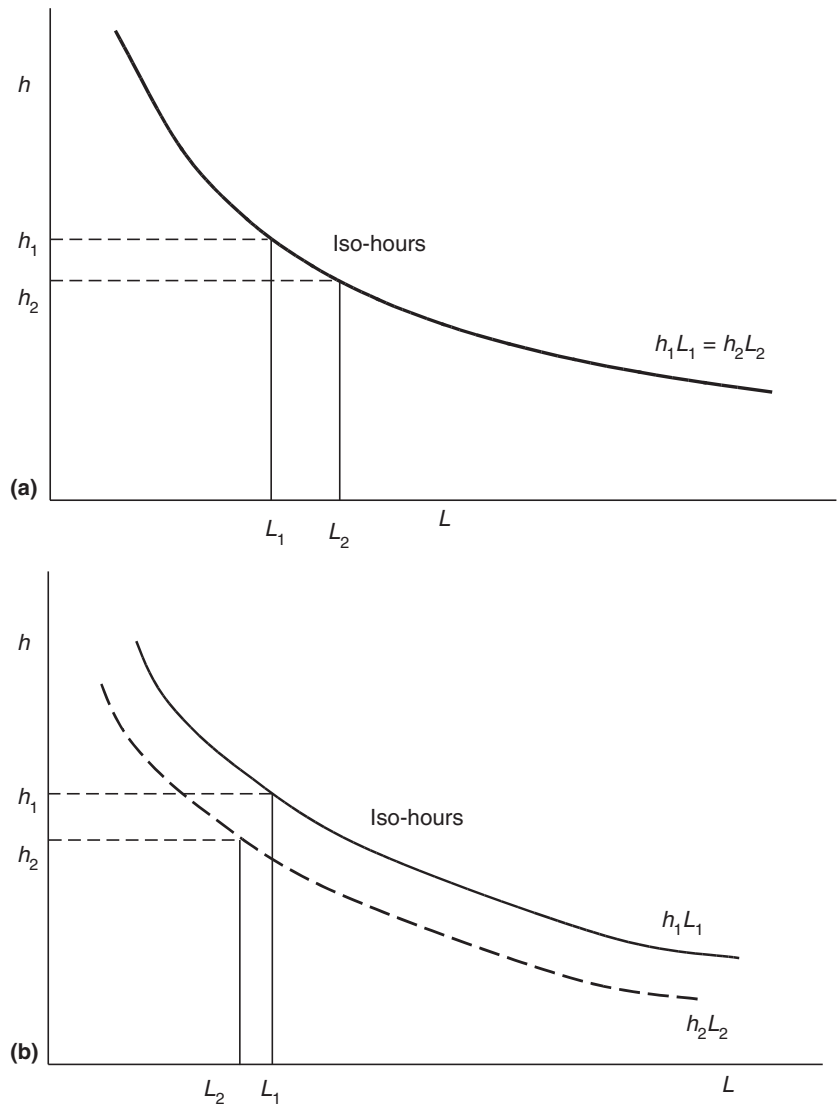


Figure 5.1 Changing Standard Working Hours ($h_1 \rightarrow h_2$): Lump of Labor (a, Upper Graph) and No Free Lunch (b, Lower Graph)

and number of workers is equivalent to the isohours curve. The iso-wage-cost curve with constant wages is depicted in figure 5.2 as the curve ABC . Figure 5.2 also illustrates the effects of overtime payments on the isocost curve. If h_1 is normal working hours and there are overtime payments, there is a kink in the iso-wage-cost curve at point B . Now the iso-wage-cost curve is the curve ABD .

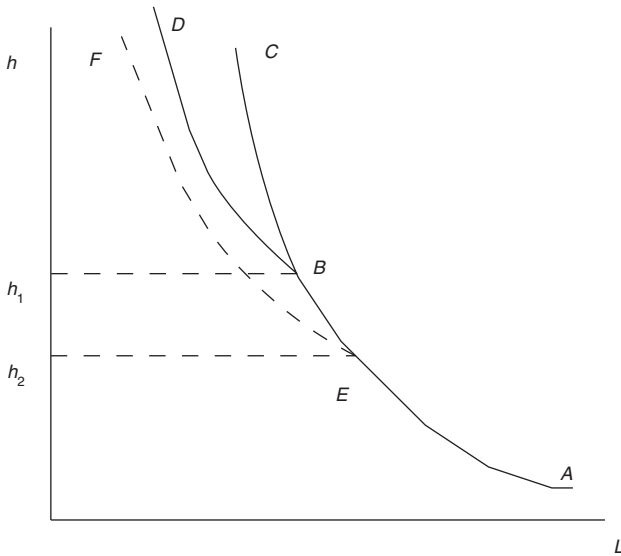


Figure 5.2 Iso-Wage-Cost Curves with Overtime Work

The kink in the iso-wage-cost curve simply implies that for the same wage cost the firm gets fewer working hours if workers work overtime. If standard working hours are reduced, the kink shifts from B to E , and the iso-wage-cost curve becomes AEF .

Calmfors and Hoel (1988) show that the employment effects of shorter working hours depend on the situation before the hours reduction. If initially workers already worked overtime, the hours reduction causes employment in terms of the number of workers to fall because the hours reduction has made the employment of a worker more expensive, while the price of marginal hours has not changed. Therefore, firms will reduce the production factor that became more expensive. If initially all workers are occupied at standard hours, the employment effects depend on the new situation. If it is optimal that the new standard hours are adopted, the number of hours will go down, and additional workers will be hired. If it is optimal that workers are employed overtime in the new situation, the employment effects are unclear.

Another main determinant of the effects of working-time reduction is the hourly wage. If working time is reduced, workers often bargain an hourly wage increase to keep the weekly wage constant. This will further decrease the demand for workers. All in all, reductions in standard working hours will only lead to an expansion of employment in very special circumstances that are not very likely to occur.

In a competitive labor market it may be that workers want to organize a reduction in working hours because this would increase their marginal product of labor

(Marimon and Zilibotti 2000). The mandatory reduction of working hours would give the workers market power so they could increase their wage. Of course, individual workers would like to deviate from the agreement and work longer hours at the higher wage, just as producers have an incentive to deviate from a cartel agreement.

Another reason for a mandatory reduction of working hours arises when employers have monopsony power. Manning (2004) argues that in a monopsony not only is the wage rate less than the value of the marginal product, but also the firm can induce employees to work more than would be optimal for the worker given the monopsony wage. In the same way that a minimum wage can be welfare improving in case of a monopsony, working-hours reduction can be welfare improving.

5.2.2 Part-Time Work

The decision about part-time work is determined by labor supply and labor demand behavior. As indicated earlier, part-time work provides flexibility to the worker that allows that worker to combine a paid job with care activities within the household. If so, labor supply will increase.

The line of reasoning is as follows. If a worker can choose only between a full-time job or no job, it may be that the reservation wage is too high for the worker to enter the labor market. This is shown in the upper part of figure 5.3. The horizontal axis from left to right gives the hours of leisure, and from right to left the hours of work. The vertical axis indicates the total income, which partly consists of labor income. A worker who works zero hours—a nonparticipant—is at point *E*, the endowment point where utility equals U_{np} . This worker could also work full-time and reach point *A* on the budget line *EA*—of which the slope represents the wage—where utility equals U_{ft} . If the labor market offers only full-time jobs, the worker will prefer not to work since $U_{ft} < U_{np}$. For this individual to work full-time, the wage would have to increase until the slope of the budget line would be at least equal to *EC*. However, if a worker can choose between a full-time job, a part-time job, or no job, it may be that the wage is sufficiently high to induce the worker to enter the labor market. This is the case in the lower part of figure 5.3 since at point *B* the worker derives a higher utility from the combination of part-time work and labor income than at points *E* and *A*. This worker will decide to enter the labor market as soon as it is possible to work part-time.

The labor market effects of the introduction of part-time jobs at an aggregate level are shown in the upper part of figure 5.4. Because of the increase in labor supply, the wage falls, and although overall employment increases, the employment of full-time workers drops. Because of the fall in the wage, fewer people are willing to work full-time. However, not only labor supply will be affected by the introduction of part-time jobs. Labor demand literature distinguishes at least three

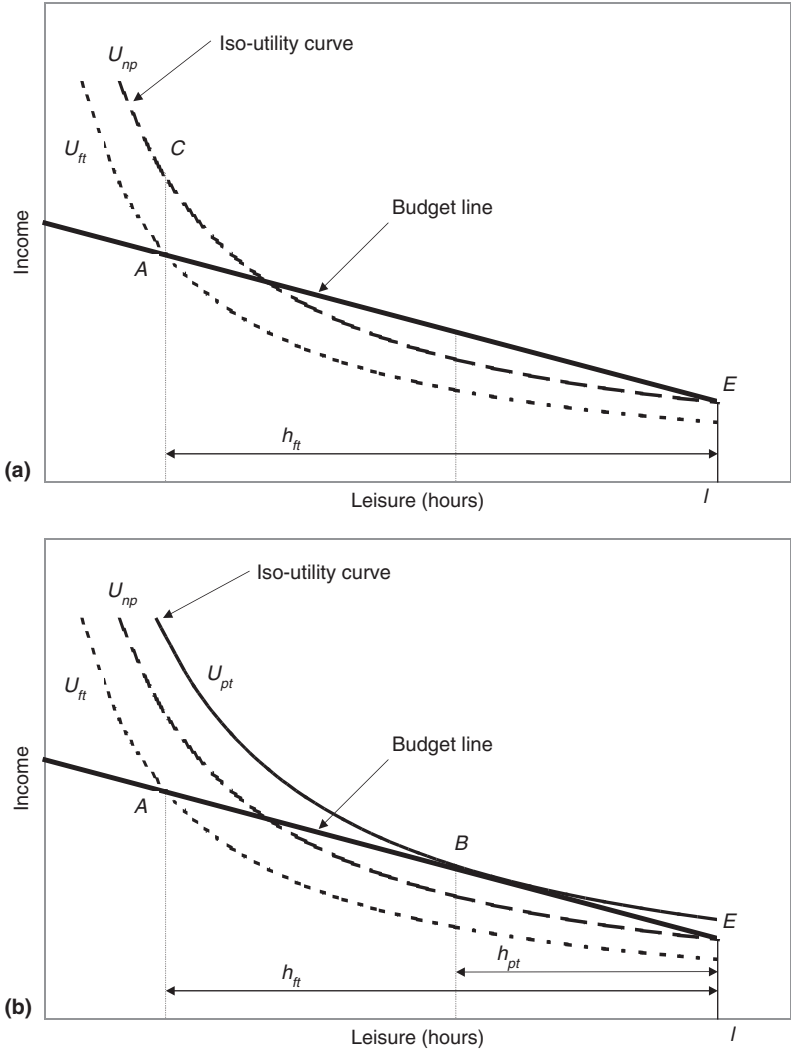


Figure 5.3 Choice of Hours of Work: Only Full-Time Jobs Available, Choice Is Nonparticipation (a, Upper Graph); Introducing Part-Time Work, Choice Is Participation (b, Lower Graph)

reasons for firms to employ part-time workers (Jaumotte 2003). The first reason concerns optimal staffing. Part-time jobs make it easier for firms to match workers to a changing workload because of demand peaks during the day or during the week. The second reason is to have cheap and flexible workers; part-time workers have lower hourly wages and less employment protection. The third reason is related to vacancies. Some vacancies cannot be filled or some workers cannot be

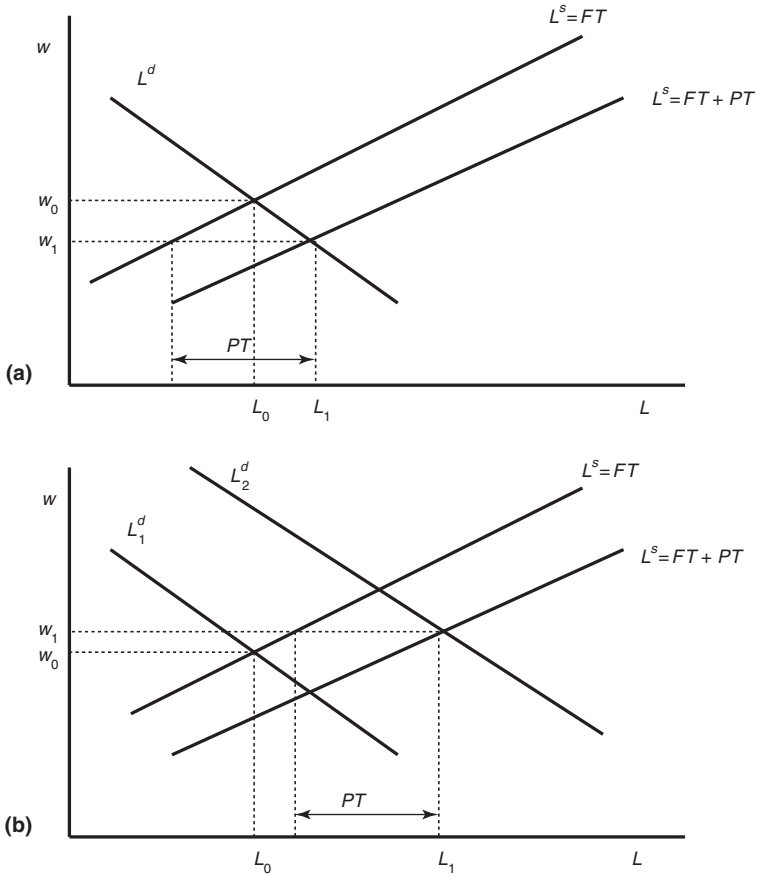


Figure 5.4 Introducing Part-Time Jobs: At the Expense of Full-Time Jobs (a, Upper Graph); Stimulating Full-Time Jobs (b, Lower Graph)

retained if no part-time jobs are offered. If employers can produce more cheaply, labor demand will expand, and the net wage effect may be negative or positive. The lower part of figure 5.4 shows a situation where the increase in labor demand causes wages to rise a bit and employment to expand a lot.

5.3 Empirical Evidence

5.3.1 Working Hours

Table 5.2 presents a long-term perspective on the annual numbers of hours worked by employees. Clearly, for all the countries presented there was a substantial decline in the annual hours of work between 1955 and 2005. In 1955 workers

in Iceland had the most annual working hours (2,391), while Italian workers had the fewest (1,958). In 2005 Korean workers still worked 2,392 hours per year, whereas working hours in Iceland had declined to 1,794 per year. Norwegian workers had the fewest, with 1,360 annual hours of work. Over the 50-year period the decline in annual working hours was strongest in Luxembourg; on average each year workers worked 17.8 fewer hours than the year before. The decline in annual working hours was smallest in Austria and New Zealand, where, on average, each year workers worked 4.6 fewer hours than the year before.

Annual hours of work are influenced by the number of working hours per week and the number of workweeks per year. Table 5.2 also presents the anatomy of a typical workweek for dependent employees in 2002. There is a wide range in weekly hours, from 30.1 in the Netherlands to 43.2 in Iceland. The low number of weekly hours in the Netherlands has to do with the high share of part-time employment (discussed later). Similarly, there is also a wide range in the number of workweeks per year, from a low of 35.4 weeks in Sweden to a high of 44.6 weeks in Greece. Most of the countries are in the range between 1,400 and 1,700 hours. On the high end are Greece and Poland with 1,816 and 1,817 respectively. The annual number of working hours is low in the Netherlands (1,223), Norway (1,339), and Sweden (1,349).

In some countries the reduction of working hours was mandatory. Through collective agreements in Germany the standard workweek was reduced from 40 to 36 hours between 1984 and 1994 (see Hunt 1999). In 1985 (West) German unions began to reduce standard hours on an industry-by-industry basis in an attempt to reduce unemployment. In the late 1970s and early 1980s unions went on strike for shorter working hours. In 1984 the unions were successful, and the standard working week in the metalworking sector was reduced from 40 to 38.5 in 1985. In exchange for shorter working hours, the unions agreed to greater flexibility in the use of standard hours, which could vary daily within a week, vary from week to week, or vary across employees. The restriction was that averaged over a period of time and across workers, the working week would consist of 38.5 hours. In her analysis to establish the extent to which work sharing affected employment, Hunt (1999) used individual-level data from the German Socio-economic Panel and industry-level samples roughly covering the period 1984–94. The main results concerning the effects of a reduction of the standard working week by 1 hour are the following. Actual working hours declined by 0.9 hour, hourly wages increased by 2.2 percent, employment of women increased by 0.2 percent, and employment of men decreased by 1.2 percent. Therefore, actual working hours follow standard hours closely. Monthly wages were hardly affected by the reduction in working hours; workers bargained sufficient increases in their hourly wage to compensate for the reduction in working hours. Finally, the results suggest that reductions in standard hours caused employment losses among men.

Table 5.2 Average Annual Hours Actually Worked per Person in Employment (1955–2005) and the Anatomy of a Typical Working Year for Dependent Employees (2002)

	Average annual hours		Average annual change	Anatomy of annual hours, 2002		
	1955	2005		Hours per week	Weeks per year	Hours per year
Australia	1,984	1,753	−4.6	—	—	—
Austria	2,086	1,519	−11.4	38.4	39	1,497
Belgium	2,346	1,611	−14.7	36.3	40	1,451
Canada	2,062	1,768	−5.9	—	—	—
Czech Republic	—	1,940	—	41.3	41	1,692
Denmark	2,070	1,552	−10.4	36.3	38.9	1,410
Finland	2,048	1,714	−6.7	38.8	38.5	1,491
France	2,040	1,434	−12.1	36.2	40.5	1,467
Germany	2,265	1,437	−16.6	36.5	40.6	1,480
Greece	2,275	1,912	−7.3	40.7	44.6	1,816
Hungary	—	1,804	—	40.9	43.9	1,798
Iceland	2,391	1,794	−11.9	43.2	39.6	1,714
Ireland	2,378	1,636	−14.8	36.3	43.7	1,585
Italy	1,958	1,592	−7.3	37.4	41	1,533
Japan	2,081	1,775	−6.1	—	—	—
Korea	—	2,392	—	—	—	—
Luxembourg	2,434	1,544	−17.8	37.9	41.7	1,582
Mexico	—	2,110	—	—	—	—
Netherlands	2,207	1,409	−16.0	31.8	38.4	1,223
New Zealand	1,980	1,749	−4.6	—	—	—
Norway	1,987	1,360	−12.6	37.3	36	1,339
Poland	—	1,994	—	41.8	43.4	1,817
Portugal	2,270	1,709	−11.2	40.4	41.8	1,688
Slovak Republic	—	1,739	—	41.8	42.2	1,761
Spain	2,041	1,774	−5.3	38.8	42.2	1,639
Sweden	1,987	1,587	−8.0	38.1	35.4	1,349
Switzerland	2,013	1,534	−9.6	37.5	42.3	1,586
Turkey	2,278	1,918	−7.2	—	—	—
United Kingdom	2,156	1,624	−10.6	38.2	40.5	1,546
United States	2,030	1,790	−4.8	—	—	—

Sources: Average annual hours: Groningen Growth and Development Center and the Conference Board, Total Economy Database, 2006 (www.ggdcc.net); anatomy of annual hours: OECD (2004b), p 35.

Notes: The estimates for average annual working hours are intended to include paid overtime and exclude paid hours that are not worked because of sickness, vacation, and holidays. Annual weeks worked = 52 minus holidays and vacation weeks, absences due to nonholiday reasons, and absences due to sickness and maternity.

In France the standard (mandatory) workweek was reduced from 40 to 39 hours in 1982 (see Crépon and Kramarz 2002 and box 5.1). Studies show that the reduction in the standard workweek also reduced the average number of hours worked weekly and increased the average wage rate, but total employment declined. In other words, work sharing held weekly incomes relatively constant but increased

unemployment. The nature of the working-hours reduction has been country specific. France had government-imposed working-hours reductions. Germany also faced weekly working-hours reductions, but they were the outcome of negotiations between unions and firms, and recently these negotiations have opened up the possibility of longer working hours (see Boeri et al. 2008). Andrews, Schank, and Simmons (2005) analyze the German labor market using establishment data to investigate the effects of working-time reduction. They find no evidence of work sharing, that is, reductions in the length of the working week leading to more jobs. This finding is irrespective of whether firms operate with workers on standard hours, working on overtime hours, or a combination of the two. Kapteyn, Kalwij, and Zaidi (2004) analyze cross-country differences in level and developments—reductions—in actual working hours that they interpret as work sharing, assuming that the reductions in actual working hours are driven by changes in standard hours.² They find that work sharing has a significant positive long-run effect on the wage rate and a positive but insignificant long-run effect on employment.

Box 5.1 Mandatory Reduction of Working Hours in France

At the beginning of 1982 the workweek in France was reduced from 40 to 39 hours. Bruno Crépon and Francis Kramarz use labor force survey data to study the effects of this working-time reduction. One way to establish the effects of working-time reduction is to compare the employment losses of two groups of workers. The first group consists of workers who worked 40 hours per week, the second group of workers who worked 39 hours. The first group was affected by the mandatory reduction of working hours; the second group was not. The employment losses cumulated over two years were as follows:

	1982–84	1985–87	Difference
40 hours	16.5	11.9	4.6
39 hours	12.6	12.1	0.5
Difference	3.9	−0.2	4.1

Whereas shortly after the introduction of the 39-hour workweek the group of workers working 40 hours per week experienced an employment loss of 16.5 percent, the group working 39 hours per week faced an employment loss of 12.6 percent, a difference of 3.9 percent. A couple of years later, in 1985, the

(continued)

² Indeed, there is ample empirical evidence that actual hours follow standard hours. The analysis of Kapteyn, Kalwij, and Zaidi (2004) is based on data from 16 OECD countries over the period 1960–2001.

Box 5.1 (continued)

difference in employment loss between the two groups was only 0.2 percent. Under a difference-in-differences approach, this result suggests that the mandatory reduction of working hours caused an employment loss of 4.1 percent for the workers working 40 hours per week. All in all, Crépon and Kramarz find an impact of annual employment losses between 2 and 4 percent. A special feature of the French situation was that hourly minimum wages were increased shortly before the workweek reduction. For minimum wage workers already employed, the nominal weekly wage was fixed, while for newly hired minimum wage workers, the weekly wage was based on the new working week of 39 hours, causing them to be 2.5 percent cheaper than workers already employed. Minimum wage workers were obviously most affected by the reduction of the workweek. Crépon and Kramarz conclude that for low-wage workers the annual probability to lose their job increased by 8.4 percentage points.

Source: Crépon and Kramarz (2002).

5.3.2 Part-Time Work

An overview of part-time employment is given in table 5.3. In 2005 part-time work among men was most common in Australia, where almost 16 percent of male employees worked part-time. Also in the Netherlands and in Japan many men worked part-time. In the Czech Republic and Slovak Republic, Hungary, and Luxembourg hardly any men worked part-time. Among women the variation in part-time work is much larger. Whereas in the Czech Republic and Slovak Republic and Hungary fewer than 10 percent of female employees had a part-time job, in the Netherlands more than 60 percent of female employees worked part-time.

In theory, part-time employment is negotiated between workers and employees. In the same way that nonstandard working hours provide flexibility to both sides of the labor market, part-time employment enables employers to match staffing and production requirements and enables employees to match working hours with family life. For employers, part-time jobs may be attractive because they also are often temporary jobs, creating much flexibility to adjust the workforce. However, this is not the case in every country. In some countries part-time jobs have all the ingredients of full-time jobs except the weekly working hours. Nevertheless, a choice for part-time work is not always voluntary for workers. Some workers work part-time involuntarily when they would prefer to have a full-time job. Table 5.3 gives an overview of the share of involuntary part-time work among part-time workers. Both for men and women there is a wide range. For men, the range is from a low of 1.8 percent in the Czech Republic to a high of 52.9 percent in France; for women, the range is from a low of 3.7 percent in the Czech

Table 5.3 Part-Time Employment (%)

	Part-time employment		Involuntary part-time		Part-timers prefer FT		Full-timers prefer PT	
	Men	Women	Men	Women	Men	Women	Men	Women
Australia	15.7	41.7	17.0	8.5	—	—	—	—
Austria	4.8	29.6	—	—	—	—	—	—
Belgium	6.2	33.1	—	—	31	25	9	36
Canada	10.8	26.9	34.9	29.8	—	—	—	—
Czech Republic	1.6	5.5	1.8	3.7	—	—	—	—
Denmark	12.0	24.9	13.1	13.9	69	8	7	21
Finland	7.9	14.8	32.8	40.2	—	—	—	—
France	5.3	23.3	52.9	38.8	69	35	11	25
Germany	7.4	39.4	17.8	12.6	52	12	5	10
Greece	3.0	11.1	50.2	36.0	33	25	8	9
Hungary	1.8	5.0	—	—	—	—	—	—
Ireland	6.8	34.8	—	—	78	30	8	12
Italy	5.3	29.2	—	—	83	42	22	32
Japan	14.2	42.3	18.9	4.0	—	—	—	—
Korea	6.5	12.5	—	—	—	—	—	—
Luxembourg	1.7	31.1	—	—	—	—	—	—
Netherlands	15.3	60.9	8.2	4.6	25	7	13	23
New Zealand	10.2	35.3	—	—	—	—	—	—
Norway	10.0	32.9	17.2	15.2	—	—	—	—
Poland	7.1	17.4	—	—	—	—	—	—
Portugal	5.9	14.4	16.1	24.1	100	40	7	2
Slovak Republic	1.4	4.1	—	—	—	—	—	—
Spain	4.2	22.2	—	—	36	37	8	14
Sweden	—	—	34.7	31.3	—	—	—	—
Switzerland	8.0	45.8	8.4	5.8	—	—	—	—
Turkey	3.2	13.4	—	—	—	—	—	—
United Kingdom	10.0	39.3	23.8	9.5	72	22	3	9
United States	7.8	18.3	7.4	8.0	—	—	—	—

Sources: OECD (1999, 2006a).

Notes: Part-time employment refers to persons who usually work less than 30 hours per week in their main job. It is measured as a percentage of total employment (2005); involuntary part-time employment is the percentage of the interviewees who state that they are working part-time because they could not find a full-time job (1997); part-timers who prefer to work full-time and full-timers who prefer to work part-time (1994).

Republic to a high of 40.2 percent in Finland. Table 5.3 also indicates to what extent full-time work and part-time work are in line with personal preferences. To the extent that individuals working part-time prefer to work full-time, there is underemployment; to the extent that individuals working full-time prefer to work part-time, there is overemployment. Here too the cross-country differences are quite large. In Portugal all men working part-time prefer to work full-time,

while in the Netherlands only 7 percent of the women working part-time prefer to work full-time. There are also large gender differences in the share of individuals working full-time who prefer to work part-time.

There is a clear negative relationship between the incidence of part-time work for females and the share of females who work part-time involuntarily. In Finland, in 1997, where only about 8 percent of the females worked part-time, 40 percent of them worked part-time involuntarily. In the Netherlands, where 61 percent worked part-time, only about 5 percent worked part-time involuntarily. In Australia 40 percent of the females employed worked part-time, and about 9 percent of them did so involuntarily. Apparently, in a situation where there are many part-time workers, part-time work is more attractive. This may have to do with unions being more interested in representing part-time workers once their number is large enough. Or it may be that part-time work can only grow if it is sufficiently attractive for workers. Nevertheless, there is concern that part-time jobs marginalize women in the labor market when part-time jobs are characterized by poor wages and benefits, low job tenure, and absence of training, reducing women's prospects of promotion and putting them at higher risk of dropping out of the labor force.

On the basis of an international overview of part-time work, the OECD (2001b) concludes that hourly earnings in part-time jobs are lower than in full-time jobs, and employer-provided training is less frequent in part-time jobs than in full-time jobs. Nevertheless, many part-time workers, especially females, have no desire to work full-time. Similarly, many full-time workers prefer to work part-time but have no opportunity to do so. Garibaldi and Mauro (2002) investigate employment growth in 21 OECD countries. They find that the increase in part-time work in many countries is not associated with net gains in total hours of work. The exception is the Netherlands, where the strong employment growth is largely accounted for by a net increase in part-time jobs in the service sector taken by women aged 25–49. The result for the Netherlands is confirmed by Van Lomwel and Van Ours (2005). They investigate the relationship between the growth of part-time work and growth of total employment in the Netherlands and find that the two developments are strongly related: the growth of part-time labor is at least partly responsible for the “Dutch employment miracle.” Manning and Petrongolo (2005) perform a cross-EU-country analysis of the pay penalty related to part-time jobs. They find evidence of this penalty, with the United Kingdom having the highest part-time pay penalty, making it difficult for women to move between full-time and part-time work without occupational demotions. At the same time, they also find that part-time work in the United Kingdom carries a higher job satisfaction premium than in most other countries. Apparently, women experience part-time jobs as having compensating wage differentials. Perhaps part-time jobs allow women to combine work and care, and therefore the lower wages attached

to part-time jobs do not cause job dissatisfaction. Buddelmeyer, Mourre, and Ward (2005) study recent developments in part-time work in EU countries. They find that part-time employment is significantly affected by institutions and policy. Both the legal framework affecting part-time positions and the creation of financial incentives (subsidies and improvement of social protection) to take a part-time job are important.

5.4 Policy Issues

5.4.1 Should Governments Regulate Working Hours?

Regulation of working hours may originate from a bargaining agreement between unions and employers or through government intervention.³ If regulation is through a bargaining agreement, governments usually do not have an incentive to intervene. If such a bargaining agreement does not exist, there may be an efficiency reason for governments to intervene and impose working-hours regulations. This may be the case if employers have monopsony power. Just as a small increase in the minimum wage may have positive employment effects but a large increase may reduce employment, there may be a small range of working-time reductions that can lead to an increase in employment, whereas a large working-time reduction will be at the expense of employment. The explanation of this phenomenon is in the noncompetitiveness of the labor market. Because of monopsony or bargaining power, labor market imperfections arise that can only be balanced through government intervention (as in the case of a minimum wage; see chapter 2). However, some governments want to influence working hours on the basis of the idea that reducing working hours will increase demand for workers and thus lower unemployment. To the extent that this is the driving force, there is no urgent need for governments to interfere. One of the few issues on which economists agree is that employment is not a lump of labor that can be redistributed without costs. On the contrary, mandatory reductions of working hours are often imposed at the expense of employment.

A century ago there may have been good reasons for the government to counter the monopsony position of employers by introducing mandatory reductions in working hours, increasing overall welfare. Nowadays it is difficult to find strong arguments for government intervention. Mandatory reductions of working hours are still advocated as an instrument to reduce unemployment, possibly because this policy seems to come as a free lunch. If the lump-of-labor story held, governments

³ Here we only discuss regulation of the amount of working hours, not the timing of working hours over days or weeks. This timing is an issue, for example, for shop opening hours and concerns coordination of leisure activities across individuals (within the same family). Governmental regulations face a trade-off between economic efficiency and welfare attached to the collective enjoyment of free time (see Burda and Weil 2006).

could reduce unemployment without much effort. Because, this is not the case, there does not seem to be an important reason for governments to interfere in the labor market and regulate working hours.

5.4.2 Should Governments Stimulate Part-Time Labor?

The differences in part-time work across OECD countries are due to differences in institutional arrangements and union resistance. Part of the opposition to part-time jobs is based on the idea that part-time jobs are created at the expense of full-time jobs (see the upper part of figure 5.4). An opposite view is that the growth of part-time jobs may even stimulate total full-time employment (see the lower part of figure 5.4). Part-time jobs can be used more flexibly than full-time jobs, for example, in peak hours. The possibility to hire a worker on part-time terms may induce employers to open vacancies that they otherwise would not have opened. It may even be the case that the creation of part-time jobs triggers the creation of full-time jobs—opening a retail shop may be possible only if sufficient part-time workers can be found to work in peak hours.

The huge increase of part-time labor in the Netherlands may serve as an example. In the course of the 1980s barriers to part-time employment were slowly removed when unions and employers reached a national agreement on this. Growth of part-time labor was also stimulated by the implementation of laws that made part-time work more attractive. In 1993 the statutory exemption of jobs of less than one-third of the normal working week from application of the legal minimum wage and related social security entitlements was abolished. Currently, most taxes are neutral, and social security benefits are usually pro rata. Also since 1993 unions and employers-representatives have recommended that employers grant workers' requests to work part-time unless there are compelling business reasons for rejection. In 1995 unions and employers signed the first proper collective agreement for temporary workers.

It is possible for governments to influence the creation of part-time jobs. Creating opportunities for part-time work to facilitate the combination of work and care seems to be an excellent idea. The availability of part-time jobs widens the menu of hours of work that workers can choose from. There is no economic argument against workers working fewer hours as long as they are not forced to do so. After all, without restrictions on the choice of individuals, higher welfare may be obtained.

5.5 Why Does Regulation of Working Hours Exist?

The determination of hours of work is rarely the outcome of a market process. Market failures occur for several reasons, in particular, because preferences of workers and employers are conflicting or because of institutional restrictions. Workers'

preferences with respect to weekly working hours, for example, may vary greatly, and providing individualized contracts may be too costly for employers. Employers may prefer that working hours vary considerably over a week because of peak hours in the production process, and accommodating may be too costly for workers (if only for social reasons). There may be institutional restrictions on hours of work because unions and employers have bargained a specific number of working hours per week, have agreed on an annual number of holidays, or have provided opportunities for workers to retire early. Negotiations about working hours cannot always be left to unions if these represent the interests of their members only. Sometimes governments influence hours of work for social reasons (family life) or because they want to influence the level or composition of unemployment (early retirement schemes).

5.6 Suggestions for Further Reading

Bruno Crépon and Francis Kramarz (2002) and Jennifer Hunt (1999) provide nice empirical examples of how a reduction of standard working hours affected employment in France and Germany (see also box 5.1). Ramon Marimon and Fabrizio Zilibotti (2000) provide a theoretical exercise where they introduce regulation of working time in a general equilibrium model with search-matching frictions. There is not much economic research on part-time work. Pietro Garibaldi and Paolo Mauro (2002) provide a general analysis of employment growth that investigates the importance of part-time work. Alan Manning and Barbara Petrongolo (2005) compare job satisfaction and part-time wages across OECD countries. One of their surprising findings is that in Britain women who work part-time have higher job satisfaction despite their lower wages. The proceedings of the 2006 conference “Are Europeans Lazy? or Americans Crazy?” (Boeri et al. 2008) provide interesting information about working hours, part-time work, and the functioning of labor markets.

5.7 Review Questions

1. What is underemployment?
2. Under what conditions does work sharing lead to an increase in employment, and how plausible are these conditions?
3. Why do firms employ part-time workers instead of full-time workers?
4. How do fixed hiring costs affect the trade-off between hours and workers?
5. How does overtime work affect the trade-off between hours and workers?
6. What happens if the standard working week is reduced in a situation where workers work overtime?

7. Why are hourly earnings in part-time jobs often lower than in full-time jobs?
8. Why do overtime premiums exist?
9. Give an explanation for the cross-country negative correlation between the share of involuntary part-time work and the share of part-time jobs in total employment.

5.8 Technical Annex: Reduction of Standard Working Hours

To analyze the effects of a reduction in standard hours, two components of the cost of labor are important. First, for every worker there are fixed costs that are independent of working hours. These are mainly the costs of hiring and firing, as well as training costs. Second, every overtime hour worked above the standard is compensated at a higher rate. The total labor costs of a firm are as follows:

$$C = (wh + \omega w(h - \bar{h})d + F)L, \quad (5.1)$$

where C is the total labor costs of the firm, w is the hourly wage, h is the actual weekly working hours, ω is the hourly overtime premium, $\bar{h}$ is the standard work-week, d is a dummy variable that has a value of 1 if $h \geq \bar{h}$ and a value of 0 otherwise, F is the fixed costs, and L is the number of workers in the firm.

For simplicity we assume that wages are constant, workers are identical, and production, y , depends only on hours and workers;

$$y = L^{\eta_L} h^{\eta_h}, \quad (5.2)$$

where $\eta_h < \eta_L \leq 1$.⁴ For a given level of production, $\bar{y}$, the firm minimizes labor costs. The Lagrangian becomes

$$\Lambda = (wh + \omega w(h - \bar{h})d + F)L + \lambda(\bar{y} - L^{\eta_L} h^{\eta_h}), \quad (5.3)$$

where λ is the Lagrange multiplier. From $\frac{\partial \Lambda}{\partial L} = 0$ we find that

$$wh + \omega w(h - \bar{h})d + F = \lambda \eta_L L^{\eta_L - 1} h^{\eta_h}, \quad (5.4)$$

and from $\frac{\partial \Lambda}{\partial h} = 0$ we find that

$$w(1 + \omega d)L = \lambda \eta_h L^{\eta_L} h^{\eta_h - 1}. \quad (5.5)$$

Taking the ratio of both equations, we find that

$$\frac{\eta_h}{\eta_L h} = \frac{w(1 + \omega d)}{wh(1 + \omega d) - \omega w\bar{h}d + F}. \quad (5.6)$$

⁴ It could be that η_L is equal to 1, so the production is proportional to the number of workers. Basically the results do not change because of this. Note that $\eta_h < 1$ because of diminishing returns to hours of work. If $\eta_h = 1$, there is no optimal solution. Note that in figure 5.1 the isohours curves assume that $\eta_L = \eta_h = 1$.

After some rewriting, we find the optimal number of hours,

$$h = \frac{\eta_h(F - \omega w \bar{h} d)}{(\eta_L - \eta_h)w(1 + \omega d)}, \quad (5.7)$$

and the optimal number of workers,

$$L = \bar{y}^{\frac{1}{\eta_L}} \left(\frac{\eta_h(F - \omega w \bar{h} d)}{(\eta_L - \eta_h)w(1 + \omega d)} \right)^{-\frac{\eta_h}{\eta_L}}. \quad (5.8)$$

From these two optima we can derive the following relationships:

1. $\frac{\partial h}{\partial \bar{y}} = 0$, $\frac{\partial L}{\partial \bar{y}} > 0$; if the scale of production increases, working hours are unaffected and employment increases;
2. $\frac{\partial h}{\partial F} > 0$, $\frac{\partial L}{\partial F} < 0$; if the fixed costs of employment increase, the hours per worker increase and the number of workers goes down;
3. $\frac{\partial h}{\partial \bar{h}} < 0$, $\frac{\partial L}{\partial \bar{h}} > 0$; if the standard number of working hours goes down, the hours per worker increase and employment goes down.

This last result indicates that reducing standard working hours reduces employment. A result that is somewhat counterintuitive is that in a firm with overtime hours, reductions of standard hours increase hours of work.

If a firm has its workers working below standard working hours ($h < \bar{h}_1$), then a reduction in standard hours to $\bar{h}_2$ may not affect the choice of the firm if it is still the case that $h < \bar{h}_2$. If $\bar{h}_2 < h \leq \bar{h}_1$ the firm will reduce working hours, and it may be that employment increases.⁵ If the working-hours reduction affects productivity, wages, and production, the employment effects depend on the magnitude of the changes in wage costs and production. If only wages increase—without productivity increase or, worse, with a productivity decline—in addition to demand for labor going down, the incentives to substitute workers by hours will increase even further.

⁵ Note that without overtime hours $\frac{\partial C}{\partial h} = wL$, while with overtime hours $\frac{\partial C}{\partial h} = (1 + \omega)wL$. If standard hours of work are reduced to a level below actual working hours, the marginal cost of an additional hour of work increases. This induces firms to reduce hours of work.

6

Retirement Programs

From a historical point of view, large-scale retirement of workers is a rather recent phenomenon. Until the beginning of the twentieth century, not many workers retired. They worked as long as they could, and if they stopped working, retirement often involved a few years of dependence on children at the end of life. In the course of the twentieth century retirement programs were introduced in industrialized countries. Today, retirement is typically an extended period of self-financed independence and leisure, and in many countries workers are even forced to retire because there is a mandatory retirement age. In many countries, because of the presence of early retirement programs, workers even retire long before the mandatory retirement age, so there is a gap between the statutory and the effective retirement age. This gap arises for a variety of reasons: the existence of early retirement programs, professions in which the official retirement age is lower than the standard age (e.g., civil servants, teachers, police, the armed forces, and firemen), and the use of unemployment insurance and disability programs by those who may not be incapable of work (Pestieau 2003). Many studies indicate that older workers strongly respond to incentives to retire early. Gruber and Wise (1997) conclude, on the basis of an analysis of 11 industrialized countries, that there is a strong correspondence between the age at which retirement benefits are available and departure from the labor force. Retirement decisions are a typical example of individuals optimizing on the basis of constraints provided by the institutional structure. In many countries the labor market attachment of older workers is low. Early retirement is an important exit route for older workers to leave work, but it is not the only one. Older workers may become disabled or sick, or they may become unemployed. Sometimes unemployment benefits (Finland and Spain), disability benefits (Australia, Austria, Ireland, Netherlands, Norway, the United Kingdom), and long-term sickness benefits (Czech Republic, Norway, Sweden) are used as pathways to early retirement (OECD 2006d, 61). In some countries older unemployed workers face less demanding search requirements or are even exempted from looking for work.

In many countries retirement is financed through a pay-as-you-go (PAYG) system where current workers are taxed to pay for current retirees. The PAYG systems are not well adapted to future demographics because in many countries the population is aging. Pension programs are either defined benefit (DB) or defined contribution (DC). DB systems typically define the final pension as a function of the length of employment attachment and final earnings. Other than the risk of becoming unemployed, which reduces the length of employment attachment, there is no risk for the worker concerning the replacement of earnings provided by the pension benefits; this replacement level is guaranteed. In DC systems the contributions to a pension fund are defined, but the level of retirement benefits depends on lifetime contributions. DC systems typically require that the entitlements acquired at retirement be transformed into annual payments. This annuitization also involves some risk because it is related to the current level of interest rates. Another factor affecting annuitization is the year of retirement: later retirement leads to higher annuity rates.

The retirement programs institution has clear relationships with other institutions. Older workers sometimes use unemployment benefits as a pre-early retirement system. Therefore, there is an interaction between early retirement and unemployment benefits (chapter 11). As will be discussed in more detail later, if older workers lose their job and become unemployed, it is difficult for them to find a job. This may be related to employment protection legislation (chapter 10).

6.1 Measures and Cross-Country Comparisons

The pension wealth of a worker is the present value of the stream of expected pension benefits. Increasing the age of retirement increases the pension wealth. The difference between the pension wealth at retirement age r and at retirement age $r + 1$ is the benefit accrual. The ratio of the accrual to the net wage earnings is an implicit *tax* on earnings (if the accrual is negative) or an implicit *subsidy* to earnings (if the accrual is positive). A negative accrual stimulates retirement, while a positive accrual stimulates continuation of work.

Table 6.1 shows that over past decades in many countries the standard age, as well as the earliest age of entitlement to public old-age pensions, has gone down, while in many other countries this age did not change at all. Exceptions to the decline of the early entitlement age are Italy (where it was the lowest to start with) and New Zealand, where this age went up. Also concerning the increase of the standard age of entitlement to public old-age pensions, Italy is an outlier. Only in a few other countries has this happened: for men and women in Greece, and for women in Australia, Belgium, Spain, and Switzerland. The most common standard age of entitlement to public old-age pensions is 65. For men there is

Table 6.1 Standard Age and Earliest Age of Entitlement to Public Old-Age Pensions; Pension Replacement Rate

	Retirement ages						Pension replacement rate	
	Earliest Males		Standard Males		Standard Females			
	1969	2003	1969	2003	1969	2003	Males	Females
Australia	65	55	65	65	60	62.5	56.4	56.4
Austria	65	65	65	65	60	60	90.9	90.9
Belgium	60	60	65	65	60	63	63.0	63.0
Canada	66	60	66	65	66	65	57.4	57.4
Czech Republic	–	58.5	–	61.5	–	59.5	64.4	64.4
Denmark	67	65	67	65	67	65	86.7	86.7
Finland	65	62	65	65	65	65	68.8	68.8
France	60	60	65	60	65	60	63.1	63.1
Germany	65	63	65	65	65	65	58.0	58.0
Greece	60	60	60	65	55	65	110.1	110.1
Hungary	–	62	–	62	–	62	102.2	102.2
Iceland	67	65	67	67	–	67	84.2	84.2
Ireland	70	65	67	66	70	66	38.5	38.5
Italy	55	57	60	65	55	65	77.9	63.4
Japan	60	60	65	65	65	65	39.2	39.2
Korea	–	55	–	60	–	60	71.8	71.8
Luxembourg	62	60	65	65	62	65	96.2	96.2
Mexico	–	65	–	65	–	65	38.3	31.7
Netherlands	65	60	65	65	65	65	96.8	96.8
New Zealand	60	65	65	65	65	65	41.7	41.7
Norway	70	67	70	67	70	67	69.3	69.3
Poland	–	65	–	65	–	60	74.9	55.2
Portugal	65	55	65	65	65	65	69.2	69.2
Slovak Republic	–	60	–	60	–	57	72.9	72.9
Spain	65	60	65	65	55	65	84.5	84.5
Sweden	63	61	67	65	67	65	64.0	65.0
Switzerland	65	63	65	65	62	63	64.3	64.3
Turkey	60	60	65	60	55	55	104.0	104.0
United Kingdom	65	65	65	65	60	60	41.1	41.1
United States	62	62	65	65	65	65	52.4	52.4

Sources: Duval (2003); OECD (2007).

Note: Pension replacement rates concern net replacement rates for mandatory retirement programs of full-career workers with average earnings.

not much variation. In 2003 the standard retirement age was highest in Norway (67) and lowest in France, Korea, Slovak Republic, and Turkey (60). Clearly, low participation rates are not driven by changes in the standard retirement age, but the reduction in the earliest possible retirement age may have been important. The incentives to retire vary greatly across countries. The net replacement rate for

public old-age pensions (mandatory retirement programs) for full-career workers on average earnings is more than 100 percent in Greece, Hungary, and Turkey, but less than 40 percent in Ireland, Japan, and Mexico.

6.2 Theory

In the theoretical literature on optimal retirement, the concept of *option value* is frequently used. Pension plans often provide a large bonus if the worker works until a certain age, typically the early retirement age. If the worker retires before that age, the option of a later bonus is lost. Continuing to work preserves the option, hence the terminology the option value of work. The main characteristic of the option value model is that a person will continue to work at any age if the expected present value of continuing work is greater than the expected present value of immediate retirement (Stock and Wise 1990). A person will postpone retirement if the option value of work is positive and will retire as soon as this option value becomes negative. In other words, a worker will retire when the value of continuing work falls below the value of retiring. The option value of working will decrease as the age of the individual comes closer to the optimal retirement age. The decline in option value does not have to be smooth. For instance, if an individual becomes eligible for an early retirement option at a particular age, there is a negative shift in the option value of working at that age. In making a retirement decision a person takes future income from work and retirement into account (see technical annex 6.8 for details on the option value model).

The determinants of the decision to retire consist of push and pull factors (OECD 2006d, 53). The push factors restrict the available set of attractive job opportunities open to older workers and consist of the true or perceived imbalance between wage and productivity for older workers, problems that firms face in adjusting their labor force due to employment protection legislation, lack of incentives to invest in human capital in order to avoid skills obsolescence, bad health, and constraints on changing working hours. The pull factors primarily consist of financial incentives that make it attractive for workers to retire. These financial incentives concern both old-age pensions and early retirement programs.

Retirement decisions are the outcome of individual optimization in the context of an institutional framework. The effect of the type of pensions on retirement age is rather straightforward. In a DB system an individual who continues to work adds to the level of pension benefits in two ways. First, the length of labor market attachment increases. Second, if the wage increases with labor market tenure, the end salary increases. The increase in future pension benefits induces individuals to stay in the labor market. As soon as an individual reaches the standard or

early retirement age, there is often no longer an increase in the expected pension benefits, while if an individual keeps on working, the number of pension years is reduced. This provides strong incentives to retire.¹ In DC systems it is also the case that continuation of work leads to higher benefits because there is an additional contribution to the pension fund and the annuity rate increases. All in all, it is not a priori clear whether there is a difference in average retirement age between the two pension systems (Banks and Smith 2006). It depends on specific design features of either DB or DC systems.

6.3 Empirical Evidence

6.3.1 Age and Employment

Table 6.2 shows the effects of the introduction over past decades of early retirement programs; the average age of transition to inactivity went down considerably. Between 1970 and 2002 this age went down as much as 14.8 years for Turkish men, but also in countries like Hungary and Poland the decline in retirement age for men was substantial. For women the largest decline occurred in Hungary and Poland. Only in Korea and, for women, in Italy was there an increase in the average age of transition to inactivity.

Figure 6.1 shows 2005 employment rates for prime-age (aged 25 to 54) and elderly (aged 55 to 64) women and men. Concerning prime-age workers, a couple of stylized facts arise from these graphs. First, the range in employment rates is much smaller for men than for women. For men, the extremes are Poland with 75 percent and Iceland with 95 percent; for women, the extremes are Turkey with 25 percent and Iceland with 85 percent. Second, in every country employment rates are higher for men than for women. The differences are smallest for the Scandinavian countries, which have high rates for women, and largest for southern-European countries, where female employment rates are relatively low. For men, there does not seem to be a strong correlation between the employment rates of prime-age and elderly workers. Apparently early retirement policies differ considerably from country to country. Indeed, the cross-country range in employment rates for men is substantial, from a low of 35 percent in Poland and Luxembourg to a high of 90 percent in Iceland. For women, there is also variation in early retirement regimes, but there is also a clear positive correlation

¹ Note that as long as the level of pension benefits depends on the final salary, workers have little incentive to reduce their number of weekly working hours shortly before retirement.

Table 6.2 Average Age of Transition to Inactivity

	Males			Females		
	1967	2002	Change	1967	2002	Change
Australia	67.3	63.2	− 4.1	67.7	60.6	− 7.1
Austria		59.6	−	−	−	−
Belgium	64.2	58.5	− 5.7	63.5	56.8	− 6.7
Canada	66.0	63.1	− 2.9	61.4	61.4	− 5.0
Denmark	−	65.3	−	−	62.1	−
Finland	69.3	60.8	− 8.5	62.2	59.8	− 2.4
France	67.3	59.3	− 8.0	66.8	59.4	− 7.4
Germany	−	60.9	−	−	60.2	−
Greece	68.1	62.4	− 5.7	64.5	60.9	− 3.6
Hungary	70.5	57.8	− 12.7	69.5	56.0	− 13.5
Iceland	−	69.5	−	−	67.8	−
Ireland	−	65.1	−	−	66.1	−
Italy	64.3	61.2	− 3.1	59.6	60.5	0.9
Japan	72.2	69.6	− 2.6	68.9	65.7	− 3.2
Korea	67.1	68.0	0.9	65.5	66.7	1.2
Luxembourg	66.1	59.8	− 6.3	66.1	59.8	− 6.3
Netherlands	−	61.0	−	−	59.1	−
New Zealand	−	64.2	−	−	62.2	−
Norway	68.0	63.6	− 4.4	69.2	62.3	− 6.9
Poland	73.4	60.9	− 12.5	72.4	58.8	− 13.6
Portugal	−	65.8	−	−	63.5	−
Spain	−	61.6	−	−	61.3	−
Sweden	69.3	63.5	− 5.8	67.0	62.0	− 5.0
Switzerland	72.5	66.6	− 5.9	73.0	63.2	− 9.8
Turkey	77.3	62.5	− 14.8	63.1	61.8	− 1.3
United Kingdom	−	63.1	−	−	61.2	−
United States	69.9	65.0	− 4.9	68.6	62.9	− 5.7

Source: OECD estimates derived from the European and national labor force surveys.

Note: The average effective age of retirement is derived from observed changes in participation rates over a five-year period for successive cohorts of workers (by five-year age groups) aged 40 and over.

between the employment rates of prime-age and elderly workers. Older workers who want to stop working have a variety of ways to do so; they may become disabled, sick, or unemployed as a pre-early retirement way to leave employment. An example of the importance of the multivariate exit routes from work is given by Bratberg, Holmås, and Thøgersen (2004), who study the effects of a change in early retirement programs in Norway (see box 6.1 for details). When a

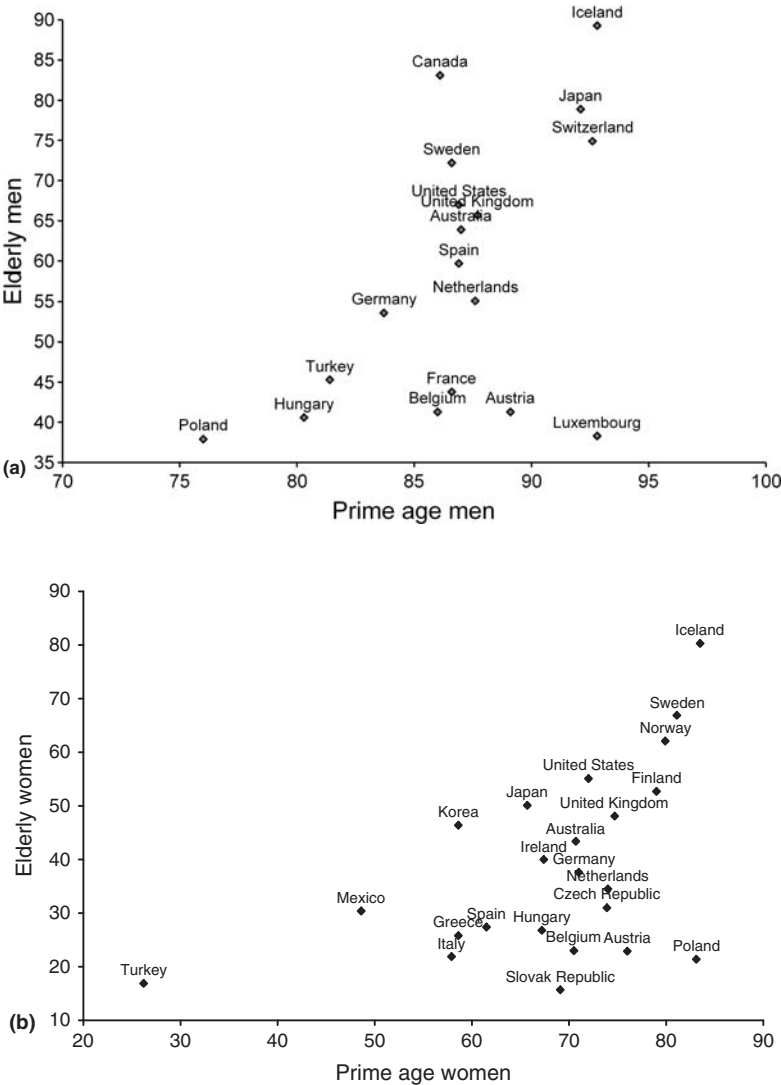


Figure 6.1 Employment Rates of (a) Men and of (b) Women, 2005

Source: OECD Labor Force Statistics.

Note: Prime age = 25–54; elderly = 55–64.

fairly generous early retirement program was introduced, many Norwegian workers who were entitled actually used it to retire early. To a large extent the early retirement system reduced the employment of older workers, and it also reduced outflow from work to other destinations such as disability, unemployment, or sickness.

Box 6.1 Stimulating Early Retirement in Norway

In Norway the standard retirement age is 67. In 1989 a fairly generous early retirement system called AFP was introduced. At the outset the entitlement age was set at 66, but this age was gradually lowered to 65 in 1990, to 64 in 1993, to 63 in 1997, and finally to 62 since 1998. In the private sector AFP is not available to everybody because only firms that participate in the central tariff agreements can offer early retirement to their employees. In these firms early retirement is possible if workers meet some rather weak conditions in terms of past work experience and past wages. The AFP system is financially attractive for many workers. At the age of 67 the worker will receive an ordinary public pension that is calculated as if that worker had continued working until age 67. Before the introduction of AFP, a significant number of older workers retired early through such unofficial routes as disability benefits, prolonged unemployment benefits, or sickness benefits. An argument put forward in favor of the AFP reform was that some workers would not stay in the labor force until 67 anyway, and the main effect of the new program was to offer them a more “dignified” exit from the workforce. Some of the changes in the system have a natural experiment character. In October 1993 the AFP age limit was reduced from 65 to 64. The first group of workers reached 64 years in the period January 1 to March 1, 1993. Espen Bratberg, Tor Helge Holmås, and Øystein Thøgersen show the labor market position of these workers three months after their birthdays.

	64 years Jan–Feb 1993			64 years Jan–Feb 1994			Diff. in diff.
	AFP firms	Non-AFP firms	Diff.	AFP firms	Non-AFP firms	Diff.	
Still at work	82.6	83.8	1.2	64.7	86.0	21.3	–20.1
AFP	–	–	–	26.0	–	–26.0	26.0
Other	17.4	16.2	–1.2	9.3	14.0	4.7	–5.9
Total	100.0	100.0	0.0	100.0	100.0	0.0	0.0

About 83 percent were still at work, while about 17 percent had a different labor market position, that is, disabled, long-term sick, or other. There was a small difference between workers in AFP firms and workers in non-AFP firms. The second group of workers reached 64 years in the period January 1 to March 1, 1994. Some of these workers became eligible for the AFP system on their birthday. As shown by the table, three months after their birthday 26 percent were in an AFP system, while about 9 percent had a different labor market position. For the workers who were not eligible for an AFP system, 86 percent were still at work three months after their birthday. Comparing both groups makes it clear that the AFP system reduced the employment rate of elderly workers significantly (20 percent) while it reduced other exits from work only slightly (about 6 percent). A Stata data file with Bratberg, Holmås, and Thøgersen’s dataset and a program (do file) are available on our webpage (see link at <http://press.princeton.edu/titles/8771.html>).

Source: Bratberg, Holmås, and Thøgersen (2004).

The low participation rates of older workers have not always been considered a problem and in fact sometimes have resulted from deliberate government policy to tackle youth unemployment or to service older workers. As the OECD (2006b) puts it: “the claim that fewer jobs for older workers result in more jobs for younger workers, though unfounded, is proving especially stubborn.” Another is that “older workers are just ‘too tired’ to carry on working.” Early retirement policies were supported by governments as a way of labor supply reduction, employers used them as a subsidized means to rejuvenate or reduce their workforce, and unions supported them because many of their older members were keen to leave the labor force. The decline in labor force participation among elderly workers is a long-term phenomenon partly stimulated by the high unemployment rates that emerged in the 1980s in many European countries. To reduce unemployment and make room for young workers, older workers often got attractive offers to retire early, “offers they could not refuse.”

Once older workers become unemployed, it takes a long time for them to find another job. Long-term unemployment among older workers is higher than among prime-age and young workers. In the literature there is not much attention to unemployment among older workers, except to the extent that unemployment is used as an alternative exit route to early retirement. Losing a job is a one-way street out of the labor force for many older workers (see Gielen and van Ours 2006 for an analysis of Dutch data). It is easier to extend ongoing firm-worker relationships than to stimulate employers to hire older workers. Nevertheless, older workers face difficulties in keeping their jobs partly because of employers’ negative perceptions about their capacities to adapt to technological and organizational changes and partly because their wage costs may rise more steeply than their productivity. Ill health and difficult working conditions, including long working hours, may also play a role. Perhaps even more important, older workers who lose their job find it extremely difficult to acquire a new job and may face large potential wage losses. On average across the OECD, the hiring rate of those over 50 is less than half the hiring rate for prime-age workers (25–49). This may have to do with employer reluctance to hire older workers but also with weak search incentives and limited support by public employment services (OECD, 2006d).

Chan and Stevens (1999) study the economic consequences of late-career job loss in the United States and find large and lasting effects on wages, assets, employment expectations, and actual employment. Chan and Stevens (2001) conclude that two years after a job loss at age 55 only 60 percent of men and 55 percent of women are employed, compared with employment rates of more than 80 percent among nondisplaced men and women who were working at the age of 55. They also find that even four years after job loss there is still a gap of about 20 percent in employment rates of displaced and nondisplaced workers. The reasons for this difference are twofold. Displaced workers return to a new job slowly, and exit rates from postdisplacement jobs are higher than from other jobs.

Daniel and Heywood (2007), referring to Hutchens (1986), argue that there are two possible explanations for the widely observed phenomenon that many firms employ older workers, but few firms hire older workers. The first explanation is related to training issues. Older workers have more experience, firm-specific skills, and knowledge and are important in training functions within firms. If they get separated from their firms, they lose this advantage and become relatively expensive as new hires. The second explanation concerns the relationship between productivity and wages. If back-loading of the compensation profile is used in order to motivate greater lifetime effort, this only holds if the worker stays in the firm. Newly hired older workers do not have an incentive to put in a lot of effort in their job, which makes them also relatively expensive. Daniel and Heywood (2007) use British data to investigate both potential explanations for the low hiring of older workers. They find that firms that use deferred compensation and internal labor markets are unlikely to hire older workers. They also conclude that modest financial incentives may be insufficient to change the hiring practices of many employers, especially employers that use back-loading and require specific human capital. Nevertheless, Behaghel, Crépon, and Sédillot (2004) find evidence that financial incentives influence hiring behavior of French firms. When a firing tax was abolished for workers hired after the age of 50, the hiring of older workers increased.

Duval (2003) presents an analysis of participation rates of older individuals that investigates the effects of a variety of incentives. He uses an unbalanced panel dataset of 22 OECD countries over the period 1967–99 containing information for three age groups, 55–59, 60–64, and 65+. The focus is on the effect of the implicit tax rates embedded in old-age pensions and early retirement programs. Changes in implicit tax rates and standard retirement ages appear to explain a major part of the trend of decline in older males' labor force participation.

Dorn and Sousa-Poza (2005) perform a cross-country analysis to study voluntary and involuntary early retirement. They conclude that generous early retirement provisions not only make it more attractive for workers to retire early but also provide an incentive to firms to encourage more workers to retire early. In particular, firms are persuaded to use early retirement programs to lay off workers during recessions and circumvent employment protection legislation.

Studies of the effects of abolishing mandatory retirement ages may be helpful to get an idea about the effects of age-related policies on aggregate labor market performance. In the United States, to abolish age discrimination, mandatory retirement before age 65 was prohibited in 1967 and before age 70 in 1978, and it was abolished altogether in 1986. Adams (2004) finds that the prohibition of mandatory retirement did not influence the probability of older workers being hired but did reduce retirement and thus increased the length of the worker-firm relationship. Neumark and Stock (1999) show that wage profiles were steeper after the abolition of mandatory retirement rather than flatter, suggesting that long-term worker-firm relationships were strengthened. They also find that the abolition of

mandatory retirement raised employment rates of older workers substantially, by as much as 7 percentage points. Von Wachter (2002) analyzes the labor market effects of the abolition of mandatory retirement in the United States. He finds that workers who were covered by mandatory retirement had a very high incidence of retirement at age 65, while those who were not covered had a much smaller incidence at age 65. The difference between the two groups gradually disappeared after the end of mandatory retirement. He estimates that the labor force of workers aged 65 or over increased by 10 to 20 percent because of the abolition of mandatory retirement.

Ashenfelter and Card (2002) find that the elimination of mandatory retirement rules for college and university professors in the United States increased their labor market attachment at higher ages substantially (see box 6.2 for details). Banks and Smith (2006), however, conclude that for the United Kingdom the employment effects of abolishing mandatory retirement are likely to be small. The main reason is that early retirement programs are more important determinants of employment rates than mandatory retirement age. They estimate that only around 1.5 percent of retirements in the United Kingdom are attributable to mandatory retirement below the age of 65.

Box 6.2 The Elimination of the Mandatory Retirement Age in the United States

College and university professors were among the few occupations to be exempted from the general elimination of mandatory retirement in the United States in 1986. Postsecondary institutions were granted a temporary exemption to enforce mandatory retirement at the age of 70. On January 1, 1994, this exemption expired. Orley Ashenfelter and David Card use a dataset of regular faculty who were aged 50 or older in the mid-1980s and were followed for the next 10–11 years to study how the elimination of mandatory retirement affected faculty retirement. The dataset concerns faculty at colleges and universities that offered defined-contribution pensions. On the basis of their analysis, the following probabilities were estimated:

	Mandatory retirement		Difference
	Yes	No	
<i>Probability to stay to age 70 (%)</i>			
From age 60	26.1	25.4	0.7
From age 65	39.2	38.6	0.6
<i>Employment outcomes if at work at age 70 (%)</i>			
Leave at 70	76.6	29.6	47.0
Still employed at 71	23.4	70.4	−47.0
Still employed at 72	8.4	51.6	−43.2
Still employed at 73	6.3	39.4	−33.1

(continued)

Box 6.2 (continued)

It appears that the elimination of mandatory retirement had no effect on retirement rates of faculty under the age of 70. However, the fraction of faculty who continued working into their 70s increased substantially. When mandatory retirement was present, of the 70-year-old faculty less than 10 percent were working at age 72. After the elimination of mandatory retirement close to 50 percent of the 70-year-olds were still working at age 72. Ashenfelter and Card also find that faculty with higher salaries or lower pension wealth were less likely to retire at any given age.

Source: Ashenfelter and Card (2002).

6.3.2 Age and Productivity

Aging may affect productivity levels for various reasons. On the one hand, older workers are thought to be more reliable and to have better skills than average workers. On the other hand, older workers have higher health care costs and lower flexibility in accepting new assignments and may be less suitable for training (Barth, McNaught, and Rizzi 1993). It is difficult to establish how age itself affects labor productivity not only because productivity is highly individual and sector specific, but also because of the interaction of age, cohort, and selection effects. Age alone is found to be a poor predictor of individual performance. There are wide variations between jobs and workers, but older workers are generally considered to be more consistent, cautious, slow, and conscientious. Furthermore, older workers have fewer accidents and are less likely to quit, thus reducing hiring costs.

The age/productivity profile is not exogenous to labor market institutions. In the past, labor market institutions have been adjusted to facilitate early departure from the labor force. Individual productivity deteriorates if no investments are made to keep human capital up-to-date. Therefore, declining productivity for older workers may be a self-fulfilling prophecy. If a worker anticipates early retirement, that worker will be less eager to invest in training to prevent productivity from deteriorating. If an employer expects a worker to retire early, that employer too will not have an incentive to invest in maintaining productivity. Current generations of older workers may have anticipated that they could retire early. If because of that they reduced investment in human capital, the concave relationship between age and productivity is caused by the existence of mandatory retirement and not the other way around.

According to Johnson (1993), most employers (and probably most employees) seem to believe in a rule of thumb that average labor productivity declines after some age between 40 and 50. This assumption is so common that few attempts have been made to gather supporting evidence; why bother to prove the obvious? Nevertheless, the variance in performance is commonly greater within age groups

than between age groups. Most research is based on cross-sectional comparisons and not on longitudinal analysis. Warr (1998) presents an extended overview of a variety of studies on the relationship between age and productivity, including studies on a range of job behaviors in conjunction with financial indicators to determine the overall outcome of employing older staff, and laboratory experiments comparing the behavior of people at different ages. According to Warr, there is no reason to think that older workers are less good at their jobs than younger workers: across jobs as a whole there is no significant difference between the job performance of older and younger workers.

Some evidence on the relationship between productivity and age is based on general data either from workers or firms. Avolio, Waldman, and McDaniel (1990) find, on the basis of a U.S. database containing individual information on personal characteristics, ability, jobs, and work performance, that the length of job experience is a better predictor of work performance than age, especially in jobs with high complexity. Hellerstein, David, and Troske (1999) use a U.S. matched worker-firm dataset to analyze the relationship between wages, productivity, and worker characteristics. They find that for prime-age workers and older workers productivity and earnings rise at the same rate over the life cycle, and they conclude that their evidence is most consistent with models in which wages rise in accordance with productivity, such as the general human capital model. Crépon, Deniau, and Pérez-Duarte (2003), using French matched worker-firm data and a similar setup of the analysis to that of Hellerstein, David, and Troske (1999), find opposite results: older workers are relatively overpaid. The age profile of wages has a concave pattern, while the age profile of productivity stops rising and even decreases after some experience level. Crépon and colleagues conclude that for workers over 35, increases in wages are not a reflection of human capital accumulation.² Ilmakunnas and Maliranta (2005) use Finnish firm data with matched average worker characteristics to investigate the relationship between wages and productivity. They find that the wage/productivity gap increases with age, which they attribute to strong seniority effects in wage setting. Dygalo and Abowd (2005) use French matched worker-firm data to compare experience/earnings and experience/productivity profiles over employment spells. They too find that earnings increase over employment spells even when productivity declines and attribute this to preferences of workers for rising earnings profiles and firms matching these preferences. However, Aubert (2003), who also uses French matched worker-firm data, does not find strong evidence of older workers being less productive than younger workers. Productivity is increasing and

² Note that the authors conclude from this that the early retirement policy in France is consistent with this evidence and that a policy of raising the mandatory retirement age may be problematic because of the poor performance of older workers in the labor market. They do not consider the possibility that the age/productivity profile may change if the attachment between workers and firm lasts longer.

concave with age. In manufacturing, construction, and trade the increase slows after age 40 and is close to zero after 50, but does not decline. In services there is also a slowdown of productivity after age 40, but even beyond this age productivity increases with age. Dostie (2006) concludes, on the basis of an analysis of Canadian linked worker-firm data, that both wage and productivity profiles are concave, but productivity diminishes faster than wages for workers aged 55 and over.³

The objective relationship between age and productivity may be difficult to establish, but employers may still have strong opinions about the productivity of older workers. Remery et al. (2003) report employers' opinions about aging issues in the Netherlands. They find that employers are less favorable—higher wage costs, lower productivity—about older workers the higher the share of older workers in the firm. About 40 percent of the employers indicated that they would not consider recruitment of older workers even if they suffered staff shortages.

6.4 Policy Issues

6.4.1 Should the Mandatory Retirement Age Be Increased?

A theoretical motivation for the existence of a mandatory retirement age is provided by Lazear (1979). According to Lazear, age/earnings profiles are upward sloping because this will discourage workers from shirking. Workers and firms engage in long-term relationships in which a worker is initially underpaid but later on in life is overpaid. Such delayed-compensation contracts will discourage the worker from shirking, but at the same time require mandatory retirement to avoid firms paying more than the value of the marginal product, averaged over the working life, that is over the duration of the contract between workers and firms. Lazear's theory requires that workers and firms want to be engaged in long-term relationships and assumes that rising earnings do not reflect increased productivity. Another complicating factor when using the age/wage profile as an indication of how age affects productivity is that the current relationship between productivity and age is at least partly determined by selective attrition from employment. The least productive workers are most likely to be the first to retire. If not accounted for, this selection effect will lead to a downward bias in the estimated productivity decline attributed to aging. Von Wachter (2002) finds that neither job tenure nor wage profiles of older workers were affected by the change in the mandatory retirement legislation in the United States. From this finding he concludes

³ Dostie indicates that it is only possible to distinguish workers by age, gender, and education. An important missing variable is occupation, so it is not possible to distinguish workers in managerial positions from workers in production positions.

that mandatory retirement is not a unique instrument to end long-term relationships between firms and workers. All in all, there does not seem to be a reason not to increase or even abolish the mandatory retirement age. But it should also be noted that increasing the mandatory retirement age may be neither necessary nor sufficient to increase labor force participation among older workers. It may not be necessary because there are other barriers that hinder older workers from keeping their attachment to the labor market. It may not be sufficient because workers use other exit routes than official retirement to leave the labor force.

6.4.2 Should Early Retirement Programs Be Phased Out?

There are basically three pension-oriented policies to reduce early retirement (OECD 2006d): reducing the generosity of early retirement benefits, improving the reward for delaying retirement, and delaying the minimum age of retirement. Most, if not all, OECD countries have reformed their pension systems using one or more of these possibilities. In addition, there may be nonpension incentives to retire early, for example, through alternative exit routes like long-term sickness, disability, or unemployment. Here too OECD countries have implemented or are implementing various policies to reduce these incentives, for example, through intensive screening of applicants for disability benefits. Finally some countries stimulate flexible pathways to retirement through part-time pensions. Although older workers may influence their retirement date, the opportunities to stay in the labor force, especially after job loss, are mainly determined by employers. OECD countries use a wide variety of policies to influence employers' behavior, including large-scale information campaigns. Furthermore, many countries, following the U.S. example, have introduced new legislation to ban age discrimination. To the extent that this legislation acts as a form of employment protection legislation, the effects are unclear. If the costs of firing older workers are raised, fewer older workers will be fired, but also fewer will be hired, making the net effect unpredictable. Another avenue of government policy to stimulate employment of older workers is the matching of costs to productivity. Many countries have introduced wage subsidy systems to hire or retain older workers. Important problems related to some of these systems are deadweight losses and substitution or displacement effects. Wage subsidies that are targeted on age alone risk being blunt instruments and may lead to small net effects at considerable costs. These age-related wage subsidies may also lead to further stigmatization of older workers.

In many countries early retirement policies were introduced as a short-term policy measure: governments wished to combat unemployment through a reduction of labor supply. When economic circumstances changed, these policies looked increasingly anachronistic (Walker 1998). Even worse, early retirement programs

may have affected employers' perceptions about the potential productivity of older workers. As long as many older workers leave the labor force well before the standard retirement age, those remaining in the labor force may be perceived as too old also. Thus older workers are trapped in a vicious circle of perceptions that lack a solid empirical basis. All in all, there is little reason not to abolish early retirement programs.

6.5 Why Do Early Retirement Programs Exist?

An important argument against the abolition of mandatory retirement is that this will create unemployment among young people. In past decades early retirement programs were introduced in several countries to reduce youth unemployment. In both cases the argument is based on the erroneous assumption that there is a lump of labor that may be redistributed from old workers to young workers without costs. Boldrin et al. (1999) show that there is no positive relationship between youth unemployment and the retirement age; that is, early retirement of older workers does not induce lower youth unemployment. Similarly, figure 6.2 shows that on a cross-country basis there is even a negative relationship between the employment rates of elderly workers (aged 55–64) and the unemployment rates of young workers (aged 20–29).

An argument in favor of early retirement programs is that older workers face health problems that prevent them from staying active in the labor market. This too seems an erroneous argument. Not only has life expectancy increased substantially over past decades, but also the health of older individuals has improved greatly. Employers may support early retirement because it makes it easier for them to get rid of older, high-paid workers. Unions may support early retirement because it is a benefit originally won through negotiation.

6.6 Suggestions for Further Reading

An overview of international differences in incentives to retire is given by Jonathan Gruber and David Wise (1997). An overview of a variety of issues related to aging populations is offered in an OECD study on aging and employment policies titled *Live Longer, Work Longer* (2006d). The main policy issue discussed is how to increase labor force attachment among older workers. The proceedings of the 2007 conference “Health, Longevity and Productivity” (Garibaldi, Oliveira, and van Ours, 2008) provide interesting information about the relationship between age, health, and productivity.

6.7 Review Questions

1. What is the difference between a defined-benefit and a defined-contribution pension system?

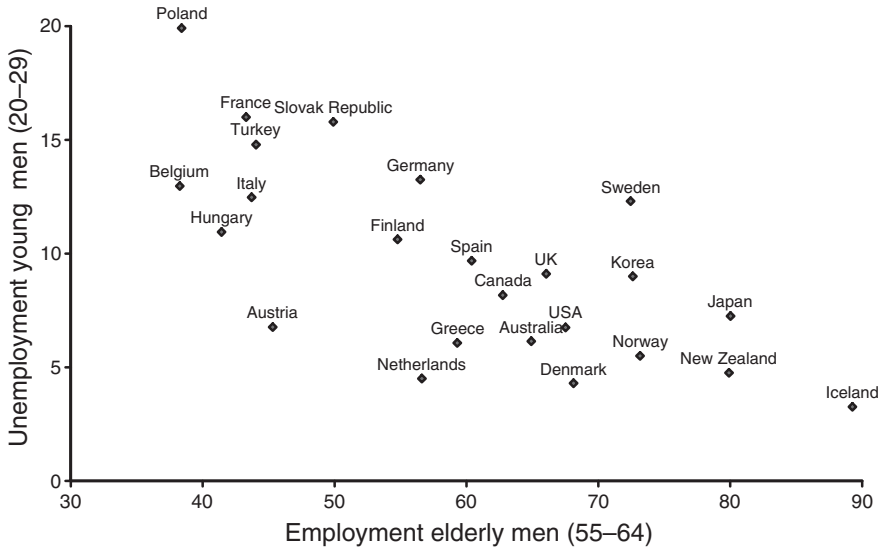


Figure 6.2 Employment Rates of Elderly Men (Age 55–64) and Unemployment Rates of Young Men (Age 20–29), 2006

Source: OECD Labor Force Statistics.

2. Why do people generally retire more gradually under a DC system than under a DB system?
3. To what extent do pension contributions differ from taxes?
4. Why is early retirement not a good instrument to reduce youth unemployment?
5. Why do wages increase with tenure, and how does that affect retirement programs?
6. How would an increase in the standard retirement age affect the behavior of employers and workers?
7. Sometimes the relationship between age and productivity is studied through the evolution of the wage over the life cycle. Why is this a dangerous strategy?
8. Explain why many firms employ older workers, but few firms hire older workers.
9. What are the essential ingredients of the option value theory of work?

6.8 Technical Annex: Optimal Retirement Age

If we assume that retirement is irreversible and that there is a constant discount rate δ , the value of retirement at a particular year r evaluated in calendar year t

consists of two components related to two subsequent periods. First, in the years from t to r the person will have annual earnings Y_t , of which the net present value can be expressed as

$$\begin{aligned} NPV_{1,t}(r) &= U_w(Y_t) + \delta U_w(Y_{t+1}) + \delta^2 U_w(Y_{t+2}) + \cdots + \delta^{r-t-1} U_w(Y_{r-1}) \\ &= \sum_{\gamma=t}^{r-1} \delta^{\gamma-t} U_w(Y_\gamma), \end{aligned} \quad (6.1)$$

where γ is an indicator of age and the U_w function indicates the indirect utility that the person derives from the wage earnings. Second, from year r until death at a given age T the person receives pension benefits $B_\gamma(r)$, of which the net present value in year t can be expressed as

$$\begin{aligned} NPV_{2,t}(r) &= \delta^{r-t} U_r(B_s(r)) + \delta^{\gamma-t+1} U_{r+1}(B_s(r+1)) + \cdots + \delta^{T-t} U_r(B_\gamma(T)) \\ &= \sum_{s=r}^T \delta^{\gamma-t} U_r(B_\gamma(r)). \end{aligned} \quad (6.2)$$

Thus the total net present value of retirement at age r can be expressed as

$$NPV_t(r) = \sum_{\gamma=t}^{r-1} \delta^{\gamma-t} U_w(Y_\gamma) + \sum_{\gamma=r}^T \delta^{\gamma-t} U_r(B_\gamma(r)). \quad (6.3)$$

Postponing retirement increases the length of the first period and reduces the length of the second period. Since earnings go up with age and pension benefits usually increase with years of service and with the wage in the last year before retirement, postponing retirement increases total income over the remaining life period. *Ceteris paribus*, this has a positive effect on the value of retirement. However, if a person postpones retirement, that person will have less time to derive utility from being retired. This second effect will decrease the value of retirement. Of the two effects, the first effect is initially more important, but eventually the second effect dominates. This means that there must be some age r^* where there is a maximum value of retirement.

Family Policies

For parents, labor supply, leisure, and child care decisions are interdependent. These decisions are determined not only by preferences and labor market conditions—wages and job opportunities—but also by family policies. Family policies consist of parental leave facilities and child care arrangements. Parental leave arrangements make it easier for mothers to stay attached to the labor market when raising children. The presence of child care affects mothers' choices with respect to leisure time versus working time. Social policies that reduce the costs to raise children—increasing the availability and quality, and reducing the costs of child care—may affect not only the labor supply decision but also fertility. These choices are affected both by the availability of formal child care and by the availability of informal child care, including family support.

To cope with aging populations, many countries face the short-run challenge of raising female labor force participation rates and the long-run challenge of raising fertility. Parental leave policies aimed at providing support for parents to combine work and care differ considerably. In some countries governments strongly support parents' caregiving, while in other countries governments rely on market-based solutions. In the United States parents rely very much on their employers to voluntarily provide options for reduced hours of work, while child care services are provided through market arrangements (Gornick and Meyers 2003). Family leave policies grant parents the right to take time off for caregiving, especially when children are below school age, and they replace some or all wages during parents' time off. Family leave policies also vary across countries on at least two dimensions: the generosity of leave available to new mothers and the degree to which policy designs encourage men's engagement in caregiving. In addition, early childhood education and care can further strengthen maternal employment by providing alternatives to full-time maternal caregiving, and high-quality early education and care can also enhance child well-being. Finally, there are working-time regulations (see also chapter 5) that can free up parents' caring time for both mothers and fathers by limiting normal hours to fewer than 40 per week and by guaranteeing a minimum number of days for annual vacations.

This chapter presents a taxonomy of policies reconciling work and family life, focusing on parental leave and child care facilities. The family policies institution has clear relationships with hours of work (chapter 5) because working-time regulations also enable parents to combine work and care. The payroll taxes institution (chapter 4) also concerns elements of family policies, in particular, income transfers to households.

7.1 Measures and Cross-Country Comparisons

Cross-country information about child care facilities and parental leave is presented in table 7.1. There are large differences in the use of formal child care arrangements for young children under the age of 3. In countries such as the Czech Republic, Austria, and Greece fewer than 5 percent of young children use formal child care arrangements, while in Denmark and the United States more than 50 percent of these children do this.¹ For children aged 3 up to mandatory school age, the use of formal child care arrangements is substantially higher. In many European countries child care use is close to 100 percent and the country with the lowest use, Greece, is still close to 50 percent. Only Japan and Korea have a substantially lower use of formal child care. Table 7.1 also shows that there are large cross-country differences in parental leave, both in duration and in level of benefits. Some countries, such as Finland and Sweden, have parental leave of about one year, while other countries, such as Australia, New Zealand, and the United States, do not have (paid) maternity leave at all. Maternity benefits range from less than 50 percent of the wage in the United Kingdom to 100 percent in many other countries. The total duration of maternity leave and child care leave ranges from 12 weeks in Mexico and Turkey to a maximum of about 3.5 years in Hungary and the Slovak Republic.

Thus all countries provide some sort of support for child care through paid leave enabling parental child care and through subsidies supporting nonparental child care, but countries differ in the way they combine these two types of support. Some countries, such as Germany, stimulate mothers to take care for children themselves by putting an emphasis on paid parental leave and providing little support for nonparental child care (see Gregg and Waldfogel 2005). The United States aims at stimulating mothers to combine work and care for young children by providing little support for parental leave and more generous support for child care. Other (Scandinavian) countries leave the choice concerning child care to parents by supporting both parental and nonparental child care.

¹ In some countries child care is arranged informally through grandmothers and relatives. In Italy, for example, informal care is the main reason for the low use of formal child care (Boeri, Del Boca, and Pissarides 2005).

Table 7.1 Summary Indicators of Formal Child Care Coverage and Maternity Leave

	Young children using formal child care (%)		Duration of maternity leave (weeks)	Maternity benefits (% of average wage)	Total duration of maternity and child care leave (weeks)
	Age < 3	Age ≥ 3			
Australia	15	60	0	0	52
Austria	4	68	16	100	112
Belgium	30	97	15	77	67
Canada	45	50	15	55	50
Czech Republic	1	85	28	69	28
Denmark	64	91	30	100	82
Finland	22	66	52	70	164
France	29	99	16	100	162
Germany	10	78	14	100	162
Greece	3	46	16	50	42
Hungary	–	–	24	100	180
Iceland	–	–	–	–	26
Ireland	38	56	14	70	42
Italy	6	95	21.5	80	65
Japan	13	34	14	60	58
Korea	7	26	8.5	100	60.5
Luxembourg	–	–	16	100	68
Mexico	–	–	12	100	12
Netherlands	6	98	16	100	68
New Zealand	45	90	0	0	52
Norway	40	80	42	100	116
Poland	–	–	18	100	122
Portugal	12	75	24.3	100	128
Slovak Republic	46	90	28	90	184
Spain	5	84	16	100	164
Sweden	48	80	64	63	85
Switzerland	–	–	16	–	16
Turkey	–	–	12	66	12
United Kingdom	34	60	18	44	44
United States	54	70	0	0	12

Source: OECD (2001a), p 144.

Note: The information about child care coverage (first two columns) concerns various years in the period 1995–2000; the category age ≥ 3 concerns children up to mandatory school age; United Kingdom is England only; the information about maternity leave concerns the period 1999–2001.

7.2 Theory

7.2.1 Child Care Facilities

The effects of child care needs and child care subsidies can be analyzed using a simple static labor supply framework in which a mother maximizes utility. The mother derives utility from consumption and leisure. If the mother takes care of the children, this is assumed to be a leisure activity. Every hour of work is at the expense of one hour of leisure, and for every hour of work, formal child care has to be provided. Informal child care is ruled out. Two extreme types of costs of child care provision are fixed costs and variable costs (see also technical annex 7.8).

Figure 7.1 illustrates how the costs of child care facilities affect individual labor supply behavior of women. In both graphs, when child care facilities are presented for free, the optimal choice of women is at point *A*, working h_A hours, where the utility of income and leisure is maximized. If there are fixed costs involved in child care—that is, costs that are independent of working hours—the budget line shifts downward, and now the maximum utility with positive working hours is at point *B*. However, assuming that in case of nonparticipation no child care is needed, the females compare utility at *B* with utility at *E* and will choose not to work. The top part of figure 7.1 can also be used to illustrate the effects of social norms and customs. In some countries there may be fixed—for example, psychological—costs related to the use of child care facilities. These costs will prevent parents from using such facilities. In economic terms the reservation wage is too low for parents to use child care facilities, and labor force participation will not occur.

The bottom part of figure 7.1 represents the situation in which child care involves variable costs; that is, the costs are proportional to the hours of work. Variable child care costs imply that the net wage is lower, so the budget line rotates. The new optimum is at *B*, and therefore working hours are reduced from h_A to h_B . In other words, because of the costs of child care, the net wage decreases, inducing women to supply less labor in the market.²

Of course, from figure 7.1 one may also conclude that providing child care subsidies may increase labor force participation of females and may increase working hours for those who are already participating.

7.2.2 Parental Leave

Parental leave is often equivalent to maternity leave, which is granted to mothers for a limited period around the time of childbirth. Some countries also have parental leave arrangements that permit both mothers and fathers additional time

² This is the case when the substitution effect dominates the income effect.

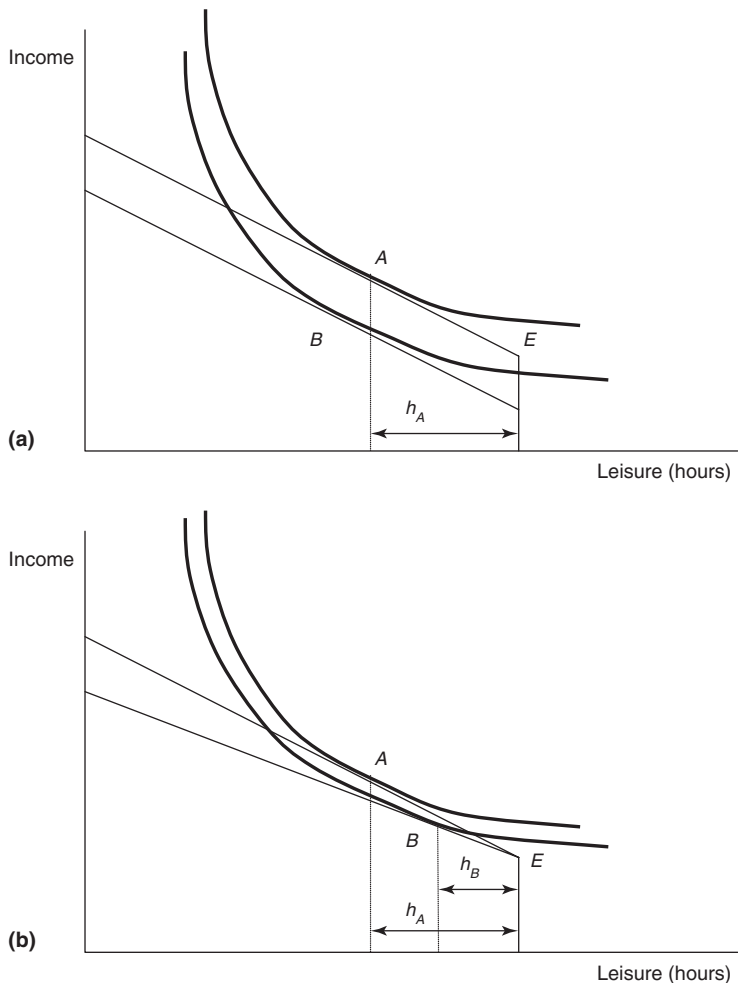


Figure 7.1 Parental Labor supply and Costs of Child Care Facilities: Fixed Costs (a, Top) and Variable Costs (b, Bottom)

off work to care for infants and young children. Basically, parental leave is a form of subsidized child care in which the subsidy is not provided to external services but to the parents who are providing child care themselves. As stated earlier, parental leave stimulates mothers to take care of children themselves while keeping links with the labor market. Like child care subsidies, parental leave will have a positive effect on labor supply of mothers, but also the demand for women will be affected, as shown in figure 7.2. Parental leave will have a positive effect on the labor supply of women because it makes participation in the labor market more attractive. However, because parental leave increases labor costs, the demand for

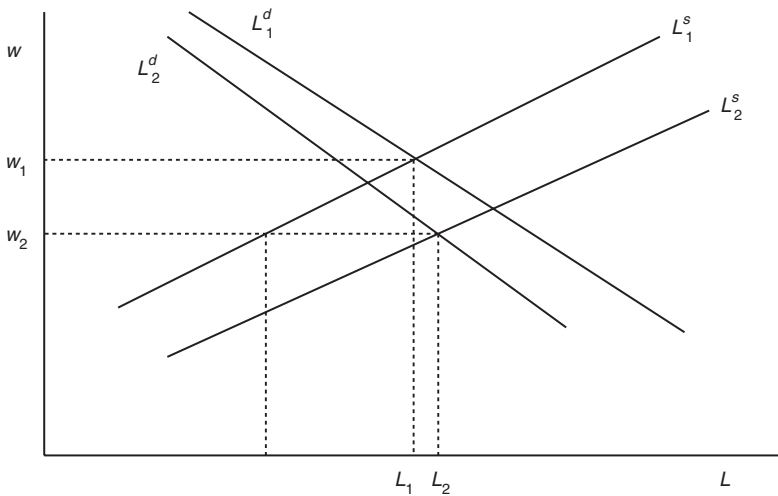


Figure 7.2 Effects of Parental Leave on Demand and Supply of Female Labor

female labor will be reduced. When these effects are taken together, there may be a small increase in employment but also a drop in female wages.

Arguments in favor of parental leave are twofold: children may be healthier if they get more parental attention, and parental leave may improve the position of women at the workplace. Arguments against parental leave are also twofold: it may reduce economic efficiency and may have particularly adverse effects on women since parental leave is most often maternity leave (Ruhm 1998).

7.3 Empirical Evidence

Female labor force participation and employment rates have increased substantially over past decades. In some countries employment rates of women are very similar to employment rates of men, while in other countries female employment rates are still very much below those of males. These differences have to do with preferences for work and family care but are also influenced by family policies. Cross-country differences themselves are determined by social and cultural customs. Family policies are also based on these customs and on the existence of market imperfections. Child care facilities, for example, are often difficult to organize at the level of individual households. A discussion of part-time work of men and women is offered in chapter 5. Here we focus on how labor market situations are influenced by family life, which has two dimensions: the presence and number of children, and child care responsibilities in the context of the family.

Table 7.2 Women's Employment Rates and Part-Time Work by Presence of Children, 2000, and Employment Rates of Mothers, 1999

	Employment rate of women				Employment rate		Part-time employment			
	Presence of children				Mothers ^a		Presence of children			
	0	1	≥ 2	Total	Couples	Singles	0	1	≥ 2	Total
Australia	68.4	55.3	43.2	66.8	48.0	—	40.8	54.1	63.1	41.8
Austria	76.0	75.6	65.7	73.5	65.7	74.8	17.4	33.6	43.7	26.7
Belgium	65.6	71.8	69.3	67.8	71.8	46.5	29.2	34.7	46.2	34.7
Canada	76.5	74.9	68.2	74.0	70.0	—	17.0	22.9	30.7	21.4
Czech Republic	80.8	72.3	59.4	73.7	—	—	2.6	4.5	7.5	4.0
Denmark	78.5	88.1	77.2	80.5	—	—	18.5	13.3	16.2	16.6
Finland	79.2	78.5	73.5	77.6	57.7	—	7.5	8.6	13.6	9.2
France	73.5	74.1	58.8	69.6	56.8	49.0	20.0	23.7	31.8	23.7
Germany	77.3	70.4	56.3	71.1	51.4	47.9	24.0	45.3	60.2	35.2
Greece	53.1	53.9	50.3	52.6	48.4	59.3	8.4	9.7	11.2	9.2
Hungary	—	—	—	61.7	—	—	—	—	—	4.9
Iceland	89.1	89.3	80.8	87.4	—	—	—	—	—	28.4
Ireland	65.8	51.0	40.8	53.1	45.5	34.1	16.6	37.2	46.4	29.7
Italy	52.8	52.1	42.4	50.7	44.9	69.5	20.0	27.2	34.4	24.1
Japan	—	—	—	62.7	33.3	—	—	—	—	38.4
Korea	—	—	—	56.3	—	—	—	—	—	8.7
Luxembourg	68.7	65.8	50.1	63.0	46.1	72.1	19.9	32.7	48.1	29.0
Netherlands	75.3	69.9	63.3	70.9	62.3	37.5	38.3	72.6	82.7	55.9
New Zealand	80.7	66.9	58.9	70.6	—	—	20.6	37.6	50.8	32.4
Norway	82.9	83.3	78.0	81.5	—	—	24.7	33.5	41.1	31.8
Poland	—	—	—	72.0	49.5	33.3	—	—	—	15.1
Portugal	72.6	78.5	70.3	73.9	70.2	81.4	11.5	10.5	11.3	11.2
Slovak Republic	—	—	—	64.8	—	—	—	—	—	2.3
Spain	54.6	47.6	43.3	50.6	41.5	61.5	13.7	17.4	18.6	15.3
Sweden	81.9	80.6	81.8	81.7	—	—	14.6	16.7	22.2	17.9
Switzerland	84.3	75.5	65.6	76.8	—	—	34.2	58.0	66.5	47.1
United Kingdom	79.9	72.9	62.3	73.1	61.3	34.2	23.7	46.6	62.8	38.6
United States	78.6	75.6	64.7	74.1	60.6	65.7	10.1	15.8	23.6	14.6

Sources: OECD (2001a, 2002).

Notes: Women aged 25 to 54 years; part-time employment: percentage of persons working part-time in total employment; Denmark, 1998; Finland, 1997; Japan, 1999; New Zealand, 2001; Switzerland, 2001; United States, 1999. Employment rates of women with children under 6 are from 1999 except Finland, 1997; Ireland, 1998; and Australia and Japan, 2000.

^a Mothers with at least one child aged under 6.

Table 7.2 gives information about the influence of the first dimension of family life, the presence and number of children. The table shows the difference in employment rates between women without children and women with one or with two or more children. The employment rate is affected strongly by the presence of children in some, but not all, countries. In Belgium, Denmark, and Portugal the employment rate of women is independent of the number of children. In many other countries there is hardly any difference in employment rates between women

without children and women with one child. Only in case of two or more children does the employment rate drop substantially.

Table 7.2 also provides information about the second dimension of family life, child care responsibilities. The table presents the employment rates of mothers with at least one child aged under 6, distinguishing between mothers who have a spouse and lone mothers. In many countries lone mothers have a higher employment rate than mothers who are part of a couple. In other countries lone mothers have a substantially lower employment rate. The differences between the labor market position of lone mothers and married or cohabiting mothers are due to the trade-off between the need to make a living and the possibility to share some of the child care responsibilities with a partner. Finally, table 7.2 shows that the share of part-time work among female employees increases with the number of children except in Denmark and Portugal. In all other countries the employment of females is affected by the presence of children because the employment rate is lower and/or because the share of part-time work increases.

Child care subsidies affect the labor market behavior of individuals. An example of a natural experiment to study these effects is presented by Piketty (1998), who finds that the introduction of a child care allowance in France reduced the employment rate of mothers substantially (see box 7.1 for more details). Naz (2004) describes a similar type of natural experiment in Norway. Here the government introduced cash benefits for parents with young children who did not use state-subsidized child care facilities. This policy change made child care facilities relatively more expensive to use and decreased the labor supply of mothers with young children (see box 7.2 for more details).

Box 7.1 Mothers with Young Children in France

In France in 1986 the Allocation Parentale d'Éducation (APE) was introduced to help parents (predominantly mothers) raise their children. The allowance of about 40 percent of the median wage (60 percent of the net minimum wage) was paid to a mother of at least three children, one of whom was younger than 3. In 1994 the allowance was extended to families with two children, one of whom was younger than 3. Mothers were entitled to the benefit from the birth of their second child, provided that it was born from July 1, 1994, onward. Because of this requirement, more and more mothers with two children, one of whom was younger than 3, became entitled to the APE benefit as time went by. By 1997 everyone in this group was entitled to the benefit. Thomas Piketty studies the effects of the introduction of the parental education allowance as a natural experiment. The development of the employment rates within the different groups of mothers—living in a couple and aged less than 55—with young children in the period 1994–97 was as follows:

(continued)

Box 7.1 (continued)

Number of children	1 child less than 3	Entitled to APE		March 1994	March 1997	Diff. in diff.	
		Before 1994	After 1994			Diff.	
2	yes	no	yes	59	47	-12	
1	yes	no	no	62	64	+2	-14
3	yes	yes	yes	31	34	+3	-15
2	no	no	no	68	69	+1	-13

This indicates that the employment rate of mothers who became entitled to the APE dropped from 59 percent to 47 percent. Such a drop in the employment rate did not occur for mothers who were not entitled to APE, that is, mothers with one child aged less than 3 and mothers with two children but no child below the age of 3. Also for mothers with 3 children of whom at least one was aged less than 3—who were already entitled to the APE before 1994—a drop in the employment rate did not occur. The employment rate among these mothers was only 31 percent in 1994, but by 1997 it had increased to 34 percent. If we use a difference-in-differences approach, it is possible to draw the conclusion that the expansion of the APE to mothers with two children caused the employment rate of this group to drop 13 to 15 percentage points. A Stata data file with Piketty’s dataset and a program (do file) are available at the website (<http://press.princeton.edu/titles/8771.html>).

Source: Piketty (1998).

Box 7.2 Child Care and Hours of Work in Norway

Female labor force participation rates in Scandinavia are high compared with most other European countries. In Norway the labor force participation rate of married and cohabiting mothers was 75 percent if the youngest child was less than 3 years old and 83 percent if the child was from 3 to 6 years old. One reason for the high labor force participation rate of women with preschool children is high-quality public child care. The costs of public and private day-care centers are shared by the state, municipalities, and parents. Parental payments for day care vary according to local rules and may be means tested. In 1998 the Norwegian government introduced cash benefits to those parents with 1- to 3-year-old children who did not use state-subsidized day-care facilities. The amount provided was roughly equivalent to the state subsidy per child given to day-care centers. Parents who used some, but not all, day-care facilities were entitled to receive a proportionally lower cash benefit. The main reason for introducing the cash-benefit reform was to offer an

(continued)

Box 7.2 (continued)

alternative to subsidized day care to parents by giving them freedom of choice in child rearing. The presence of young children increases the value of women's time at home, which reduces their labor market activity. The cash-benefit reform increases the value of mothers' time at home, thus increasing the relative costs of child care centers for parents and making market work less attractive. To investigate the effects on hours of work of mothers and fathers, Ghazala Naz compared the labor supply of parents with children aged from 1 to 3 and the labor supply of parents with children in the age group 3 to 6. Before the reform neither group of parents was entitled to the cash benefits, while, after the reform the first group of parents was entitled while the second was not. The table indicates what happened to the average number of hours worked (per week):

	Children age 1 to 3			Children age 3 to 6			Diff. in diff.
	Before	After	Diff.	Before	After	Diff.	
Mother's working hours	24.4	23.7	-0.7	24.5	26.5	2.0	-2.7
Father's working hours	40.9	41.3	+0.4	40.8	40.8	0.0	+0.4
Total working hours	65.3	65.0	-0.3	65.3	67.3	2.0	-2.3

Hence the mother's working hours decreased by 0.7 hours per week if she had children aged 1 to 3, while it increased by 2.0 hours per week if she had children 3 to 6 years old, leaving an effect of the cash benefit of 2.7 hours. The labor supply of fathers hardly changed. A Stata data file with Naz's dataset and a program (do file) are available at the website (<http://press.princeton.edu/titles/8771.html>).

Source: Naz (2004).

Ruhm (1998) presents an empirical analysis of the economic consequences of rights to paid parental leave in nine European countries in the 1970s through the early 1990s. He argues that parental leave mandates will shift the labor supply curve of women to the right while moving the demand curve to the left (as in figure 7.2). If leave benefits are primarily state funded, the effect on the demand curve will be small. This should have a positive effect on employment of women and a negative effect on their wages. Ruhm's empirical analysis exploits changes in paid parental leave policies in the countries investigated. He finds that rights to short periods of paid parental leave increase the employment rates of women by 3 to 4 percent while having little effect on wages. He also finds that extended parental leave periods raise the female employment rate by the same amount, but they also decrease hourly earnings by around 3 percent. Tanaka (2005) presents a cross-country time-series analysis investigating the relationship between parental

leave (paid and unpaid) and child health. The main result of the study is that job-protected paid parental leave decreases infant mortality, whereas unpaid leave and non-job-protected paid-leave do not.

Fertility is assumed to be affected by family income and the price of children. *Ceteris paribus*, the number of children increases with family income. This relationship represents the pessimistic view of Thomas Malthus and led to economics being labeled the “dismal science.”³ However, income effects are only part of the story. The desired number of children also depends on their price. An increase in the wage rate of the parent most responsible for child care will make a child more expensive. The latter effect suggests that there is a negative relationship between female employment and fertility. A remarkable phenomenon in past decades is the change in the cross-country correlation between female participation rates and fertility rates (Boeri, Del Boca, and Pisarides 2005). Up to the early 1980s this correlation was negative, suggesting a trade-off in time between paid work and child care. From the late 1980s onward a positive cross-country correlation is observed: countries with higher participation rates also have higher fertility rates. This change in correlation is thought to be related to more generous parental leave, greater availability of child care, and greater opportunities for flexible hours and part-time employment. Indeed, as shown in figure 7.3, there is a positive cross-country correlation between the total fertility rate and the employment rate of prime-age females.⁴ In the Nordic countries with generous provision of child care facilities and maternity leave, high fertility rates coincide with high female employment rates. In southern European countries, where these arrangements are less generous, low fertility rates coincide with low female employment rates. To some extent, the positive correlation is driven by Italy, Greece, Poland, and Spain, which combine a low employment rate of prime-age females with a low fertility rate. Nevertheless, even if these countries are ignored, there is still a slight positive correlation between fertility and female employment rates.

³ It is often stated that Thomas Carlyle gave economics the nickname “dismal science” as a response to the writings of Malthus, who predicted that starvation would result as projected population growth exceeded the rate of increase in the food supply. However, others claim that the phrase “dismal science” first occurs in Carlyle’s 1849 tract titled “Occasional Discourse on the Negro Question,” in which he was arguing for the reintroduction of slavery as a means to regulate the labor market in the West Indies. Carlyle argued that slavery was actually morally superior to the market forces of supply and demand promoted by economists since, in his view, the freeing up of the labor market by the liberation of slaves had actually led to a moral and economic decline in the lives of the former slaves themselves (see www.wikipedia.com).

⁴ The total fertility rate is the average number of children that would be born per woman if all women lived to the end of their childbearing years and bore children according to a given fertility rate at each age.

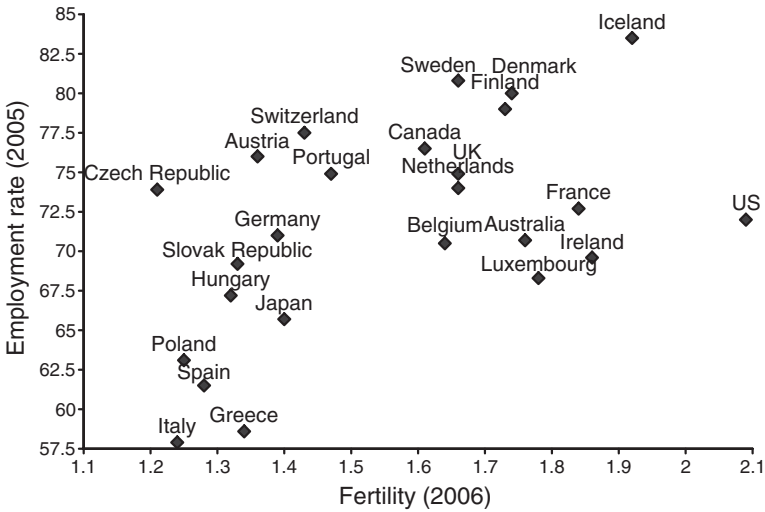


Figure 7.3 Fertility Rates and Employment Rates of Prime-Age Women

7.4 Policy Issues

7.4.1 Can Work and Family Life Be Balanced?

Governments may support external child care by subsidizing providers or consumers.⁵ Since only child care providers who meet certain requirements are eligible for subsidies, only formal providers, such as day-care centers and preschools, benefit from a policy of subsidizing providers. Consumer subsidies may be provided through child care tax credits and deductions and are usually not limited to specific types of child care providers (Blau 1992). An alternative is to provide vouchers to families with children, which could also support informal child care.

Balancing work and family life is an issue that is related to the promotion of female employment rates, but it also touches upon the quality of child care. According to Blau (2001), child care quality in the United States is low not because of a failure on the supply side of the market, but because of low willingness to pay by parents. The majority of child care subsidies in the United States do not place restrictions on the quality of the care and do not provide incentives to use high-quality care, but the subsidies are available only to employed parents. According to Blau, the child care subsidies encourage employment of both parents in two-parent families and of the single parent in one-parent families, but it is not clear why society should wish to provide this encouragement. The system of child care subsidies increases the well-being of families in which both parents are employed,

⁵ Apart from subsidies, governments may also influence child care facilities through regulations concerning child/staff ratio and the qualifications of the provider.

but does not provide benefits to families in which one parent stays home to take care of children. Indeed, as shown in box 7.2 on child care subsidies in Norway, if subsidies are provided irrespective of whether child care is actually used, many parents, mainly women, opt for providing child care themselves. If child care subsidies are provided conditional on parental employment, this will increase the demand for quantity of child care but not for quality of child care. Blau (2001) argues that there is no compelling economic or moral reason to encourage employment of both parents in a two-parent middle-class family. For single-parent families there may be a reason for child care subsidies, but this is at best an indirect approach and at worst an ineffective policy instrument to accomplish this goal. If governments subsidize the costs of raising children, they could do this without favoring market costs for child care over the forgone-earnings cost of a parent who stays home to care for a child. All in all, the main argument for subsidizing child care facilities in order to stimulate mothers to enter the labor market is based only on the idea that labor force participation should go up in order to deal with an aging society and a shrinking tax base. From the perspective of what is best for children, providing child care subsidies is not sufficient to guarantee high-quality child care arrangements.

7.4.2 Is There a Trade-off between Fertility and Employment?

Employment policies interact with cultural and social customs. In some countries the use of external child care facilities may be small because of a social stigma related to sending children to these facilities. In economic terms this implies that there are fixed—mainly psychological—costs related to the use of child care facilities. Therefore, under similar economic conditions child care facilities may be used more often in some countries than in other countries. It may also be that policies to promote female employment are not in line with social norms.

Figure 7.4 shows that there is a negative cross-country relationship between opinions about mothers working and the employment of prime-age women. In countries where many people agree with the statement that preschool children suffer when their mothers work, employment rates of prime-age women are lower. In Denmark, for example, where many disagree with this statement, female employment rates are quite high. In Italy many agree, and the female employment rates are rather low. Nevertheless, there are also clear outliers. In Sweden many think that a preschool child suffers when the mother works, but still the female employment rates are high; in Ireland and Spain many individuals do not worry, but the female employment rate is not very high. Clearly there is not a one-to-one relationship between social and cultural norms and female employment. Nevertheless, if low female employment rates in some countries represent social and cultural preferences, it may not be optimal from a welfare standpoint to promote female

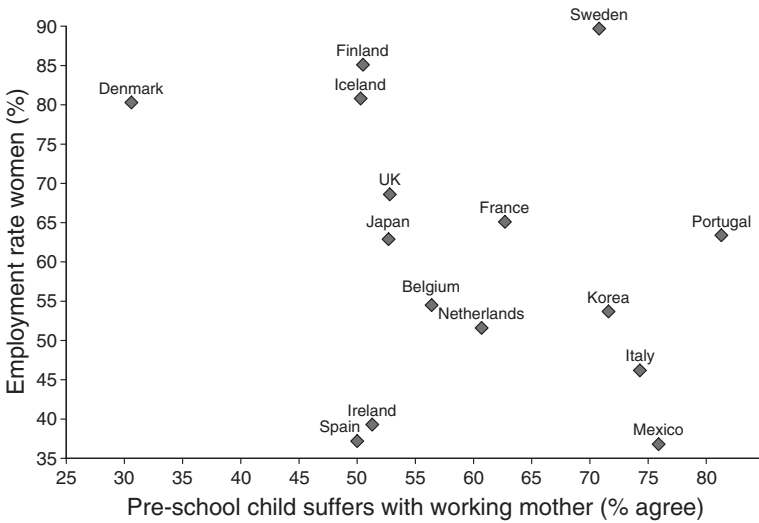


Figure 7.4 Cross-Country Relationship between Norms about Women Working and the Employment Rate of Prime-Age Women (1990)

Sources: World Value Survey and OECD (2001a).

employment.⁶ Still, it is also possible that the cultural norms merely support the existing situation. In that case an increasing female employment rate could cause opinions to change.⁷

Fertility rates are below replacement level in all OECD countries, and in some countries they are barely even half that level (OECD 2001a). In almost all OECD countries there are similar trends in female employment rates and fertility rates. Over past decades successive cohorts of young women have had higher employment rates, but lower fertility rates. Furthermore, it seems that larger increases in employment have been associated with larger decreases in fertility. In coming decades the working-age population is likely to decline in many OECD countries. Increasing female labor force participation rates may balance this decline in the working-age population. However, if increasing female labor force participation would cause lower fertility rates, that would reduce the future working-age population. At the same time, high levels of female employment are not incompatible with relatively high fertility rates. Indeed, as shown in figure 7.3, there is now even a positive cross-country correlation between fertility and employment rates

⁶ See Algan and Cahuc (2005) for more details.

⁷ In the Netherlands, for example, in 1965, 84 percent of the respondents, aged 17 to 70 years, indicated that working of married women with school-age children meets with objections. In 1970 this was 44 percent, in 1980, 36 percent, in 1991, 20 percent and in 1997, 18 percent. This change in opinions coincides with a strong increase in female labor force participation (Sociaal en Cultureel Planbureau 1998)

of prime-age women. Thus with regard to the size of the future working-age population, there does not seem to be a trade-off between female employment rates and fertility rates.

7.5 Why Do Family Policies Exist?

If two parents both have a job, an important question is who will take care of the children and who will pay for that care. Many European countries provide public and heavily subsidized child care. In a few countries there is a market for child care. Only the United States has a child care policy that is mainly private. Blau (2001) distinguishes two dimensions along which child care programs may be organized. The first dimension is the employment requirement of the program, with no requirements at the one extreme and full-time parental employment at the other extreme. The second dimension is the quality of the child care required in order to be eligible for a subsidy. According to Blau, in 1999 in the United States only one-third of the programs had a major focus on quality, while the other two-thirds had little emphasis on quality, but strong emphasis on parental employment. In European countries child care is often publicly provided with relatively generous funding and coverage and with a strong emphasis on quality. Child care subsidies may be provided because of the existence of imperfections in the market for child care. Parents have imperfect information about the quality of child care. This may lead to moral hazard and adverse selection. Moral hazard may occur because parents know less about the quality than the provider, and monitoring quality is costly. Adverse selection of providers may occur in the informal care sector. Because family day care is a low-wage occupation, women with high wage offers in other occupations are less likely to choose the day-care sector. Subsidies targeted at high-quality child care may induce parents to opt for higher-quality child care. Although this does not solve the information problem, it may induce parents to choose a level of child care that is not suboptimal from a social point of view. The externality argument is related to the same line of reasoning as that for subsidies on education (see chapter 8). High-quality child care increases the probability that children will attend higher education, which reduces costs to society related to low education: low wages, unstable employment, long-term unemployment, dependency on social transfers, and so on. Because parents may not be fully aware of these benefits, there is a difference between private and social discount rates.

7.6 Suggestions for Further Reading

The chapter “Balancing Work and Family Life: Helping Parents into Paid Employment” in OECD (2001a) is very informative about cross-country differences in

family policies. David Blau (1992, 2000) discusses child care subsidy programs, focusing on the United States.

7.7 Review Questions

1. Why do lone-parent households face above-average poverty risk?
2. Why do some countries provide family allowances only to those who are not working?
3. Why could it be welfare improving if governments subsidize child care facilities?
4. Why did the OECD cross-country correlation between female participation and fertility rates change from negative to positive?
5. How does child care affect the reservation wage?
6. How does parental leave affect employment and wages?
7. For a long time economics was called the “dismal science” because of Malthus’s ideas. What is wrong with Malthus’s ideas?
8. How can governments influence the market for child care? What are the pros and cons of these policies?
9. How could subsidies influence the choice between formal and informal child care?

7.8 Technical Annex: Child Care Facilities and Hours of Work

We assume that the mother derives utility from consumption, c , and leisure, l : $V = V(c, l)$. As indicated in the main text, if the mother takes care of the children, this is assumed to be a leisure activity. Every hour of work is at the expense of one hour of leisure, and for every hour of work, formal child care has to be provided. Informal child care is ruled out. First, we consider the fixed-costs case. This is the case if the child care facilities are provided at an entrance fee, a subscription to facilities that is unrelated to the amount of use. If w is net hourly wage, h is the (weekly) hours of work, m is nonlabor income, and F_c is the fixed cost of child care, the mother can choose between two weekly budget constraints:

$$c_1 = m \tag{7.1}$$

in case the mother decides not to work and

$$c_2 = m - F_c + wh \tag{7.2}$$

in case the mother enters the labor market and works h hours. The mother will decide to enter the labor market and work h hours if $V(c_1, l) < V(c_2, l - h)$. Even

if the earnings exceed the fixed costs of child care, the mother may not enter the labor market because of the reduction in leisure. If a child care subsidy is provided, this raises the potential income from working, causing the likelihood of work to increase.

The second case is the one where child care costs are variable; that is, every hour of child care costs c_c has to be paid. Now the budget constraint is

$$c_3 = m + (w - c_c)h; \quad (7.3)$$

in other words the budget constraint rotates because the actual net wage decreases. This may lead mothers to abstain from entering the labor market. A child care subsidy will increase the actual net wage, which increases the likelihood of work since the subsidy has only a substitution effect. If a mother is already in the labor market, a child care subsidy has an ambiguous effect on hours of work. There is a positive substitution effect and a negative income effect if we assume that leisure is a normal good.

This page intentionally left blank

Education and Training

In many countries public spending on education increased significantly at the end of the nineteenth century. In 1880, for example, the United States spent 1 percent of its GDP on primary and secondary public education. By 1920 this figure had doubled, and by 1980 it had reached 4.1 percent. From 1870 to 1970, the days attended during the school year doubled, and the fraction of those 10 to 19 years old enrolled in school increased from about 40 to 90 percent (Rangazas 2002).

Schooling and training are investments by individuals and firms where costs are paid in exchange for expected future profits. Human capital of individuals is influenced by their innate ability, their investment in schooling, and their investment in on-the-job training.¹ Going to school is a productive investment since additional years of schooling lead to higher wages and better employment prospects. On-the-job training is also a productive investment because it increases general or firm-specific productive skills of workers. Formal schooling most often takes place before an individual enters the labor market. Subject to their ability, individuals choose the educational level they wish to obtain by balancing costs and expected benefits, such as a higher wage and a higher employment rate because productivity increases with educational attainment. Training most often takes place after entrance into the labor market. Sometimes training is general; that is, it is useful in many firms. Other times training is firm specific; that is, it only has a positive effect on productivity within the firm but cannot be used in other firms. The literature on schooling focuses on the decisions of individuals concerning their choice of educational attainment; the literature on training focuses on the question of who pays for the training: the worker or the employer. Standard human capital theory predicts that the costs of general on-the-job training have to be paid by workers, while the costs of firm-specific training have to be paid by firms. An important element in the discussion is the

¹ The main difference between human capital and physical capital concerns property rights. Whereas physical capital can be sold, human capital is embodied in persons. Therefore, in order for firms to use human capital, workers and firms have to agree on the terms of use.

influence of the nature of the labor market, that is, whether it is competitive or not.

The main market failures concerning education and training are fourfold. First, individuals are confronted with incomplete capital markets, so they are restricted in their optimal decision making, and their investment decision may be suboptimal. Second, it may be that the private rate of return differs from the social rate of return. Third, there is a long time lag between an educational decision and the outcome of that decision. Fourth, there is a holdup problem: once an investment in training is made by the employer, the worker can leave the firm and negotiate a higher wage somewhere else. Alternatively, once a worker has invested in human capital, he may be fired and thus be unable to benefit from the investment. Because of these market failures, governments influence education and training, often by making it cheaper for individuals to invest in their human capital and thus increase their productivity.

There are clear relations with other labor market institutions. Payroll taxes (chapter 4) are important because educational and training decisions may be influenced by the potential net gains of those decisions. There is also an overlap with unions (chapter 3) because union-induced wage compression makes investment in education and training less profitable. Employment protection (chapter 10) may stimulate on-the-job training investments because employers want to use internal labor markets. In the absence of employment protection, workers may be reluctant to contribute to such investments.

8.1 Measures and Cross-Country Comparisons

The way formal education is organized is very country specific. Individuals may choose the amount of education they wish to acquire on the basis of costs and benefits. Acquiring formal education is an investment decision, with rates of return connected to that decision. Provision of in-company training also differs significantly across countries.

Table 8.1 provides a cross-country overview of expenditures on educational institutions as a percentage of GDP for all levels of education. There are substantial differences in spending levels, ranging from a low of 4.1 percent in Greece to a high of 7.4 percent in Iceland. Table 8.1 also gives an overview of educational attainments in terms of the average number of years spent in formal education. There are differences between men and women, but the within-country differences are much smaller than the between-country differences. Portugal has the lowest educational attainments, with men reaching 8.1 years and women 8.4 years of formal education. Educational levels are highest in Norway and the United States, where adult men, on average, have 13.8 years of formal education, while adult women, on average, have 13.9 years of formal education. There is a positive relationship between educational expenditures and educational attainment, but the

Table 8.1 Educational Expenditures and Attainments

	Educational expenditures	Years of formal education		PISA math score
		Men	Women	
Australia	6.0	13.0	12.8	108
Austria	5.7	12.1	11.5	105
Belgium	6.4	11.3	11.2	110
Canada	–	13.0	13.1	110
Czech Republic	4.4	12.5	12.3	107
Denmark	7.1	13.6	13.7	106
Finland	6.0	11.9	12.2	113
France	6.1	11.7	11.4	106
Germany	5.3	13.7	13.1	104
Greece	4.1	10.7	10.3	92
Hungary	5.6	11.8	11.5	101
Iceland	7.4	13.7	12.9	107
Ireland	4.4	12.8	13.1	104
Italy	4.9	10.2	9.9	96
Japan	4.7	12.6	12.1	111
Korea	7.1	12.4	11.3	112
Luxembourg	–	13.7	13.2	102
Mexico	6.3	8.9	8.5	80
Netherlands	5.1	13.1	12.7	111
New Zealand	6.8	12.5	12.7	108
Norway	6.9	13.8	13.9	102
Poland	6.1	11.5	11.8	101
Portugal	5.8	8.1	8.4	96
Slovak Republic	4.2	12.5	12.3	103
Spain	4.9	10.6	10.4	100
Sweden	6.9	12.3	12.6	105
Switzerland	6.2	13.4	12.3	109
Turkey	3.8	9.9	9.3	88
United Kingdom	5.9	12.8	12.6	–
United States	7.2	13.8	13.9	100

Source: OECD (2005a).

Note: Expenditure on educational institutions as a percentage of GDP for all levels of education, 2002; Italy and the Netherlands, 2001; average number of years in formal education of those 25 to 64 years old, 2003; PISA normalized score of 15-year-olds in mathematics, 2003 (United States = 100).

relationship is imperfect, with a mild slope. Portugal and Mexico are clear outliers with relatively high expenditures and low educational attainments.² Finally, table 8.1 also gives an overview of mathematics scores, as measured through the

² Of course, the comparison of expenditures and educational attainment is far from perfect since the expenditures concern a particular year, 2003, while educational attainment concerns persons aged 25 to 64 years, most of whom have been to school decades ago. Nevertheless, even among younger cohorts Portuguese men and women have relatively low educational attainment.

PISA survey.³ Here too there is large cross-country variation, with Finland reporting the highest score and Mexico having the lowest score. Again, there is a positive but mildly sloping relationship between educational expenditures and math scores. Finland has the highest math score, but far from the highest educational expenditures, while the United States has the highest educational expenditures but is among the lowest in math scores of its 15-year-olds.⁴

Table 8.2 gives an overview of the cross-country variation in employer-sponsored training programs. Obviously this type of information is not always easily available since for many countries this type of information is missing. Furthermore, there are differences in the numbers depending on whether the information comes from workers (columns 1 and 2) or from firms with at least 10 employees (columns 3 and 4).⁵ In Belgium, for example, the participation rate in employer-sponsored training programs is 13 percent according to workers and 41 percent according to firms. This difference is either due to perception or because very small firms are excluded from the firm survey.⁶ In the countries for which information is available, there is wide variation. According to information from workers, in Ireland the participation rate in employer-provided training is only 10 percent, while it is 45 percent in Norway. Also, the number of hours spent in employer-based training shows wide cross-country variation, from 8 hours per year in Italy and Poland to 36 hours per year in Denmark. According to information from firms, the variation in participation ranges from a low of 12 percent in Hungary to a high of 61 percent in Sweden. The average annual hours per worker range from 4 in Poland to 22 in Denmark.

8.2 Theory

8.2.1 Schooling

Differences between individuals' human capital may be due to differences in their innate ability, investment in schooling, and investment in nonschooling characteristics. Innate ability represents the productivity of the worker without any investment. The individual worker has to decide how much to invest in schooling.

³ PISA is a three-year survey of the knowledge and skills of 15-year-olds in the principal industrialized countries. The survey assesses how far students have acquired some of the knowledge and skills that are essential for full participation in society.

⁴ Note that this may also be due to large expenditures in tertiary education, which most 15-year-old individuals have not yet started.

⁵ In addition to these differences, firm-specific training may be underestimated because the surveys do not account for apprenticeships, for example, in Germany, where this is a very important part of firm-specific training.

⁶ There is a positive relationship between firm size and participation of workers in employer-sponsored training programs. In Belgium participation was 20 percent in firms with 10 to 50 workers, while it was 66 percent in firms with more than 1,000 employees.

Table 8.2 Cross-Country Variation in Employer-Sponsored Continuous Vocational Training Courses

	IALS data		ECVTS data	
	Participation rate	Annual volume	Participation rate	Annual volume
Australia	24	15	–	–
Austria	–	–	31	9
Belgium	13	10	41	13
Canada	28	17	–	–
Czech Republic	16	13	42	10
Denmark	45	36	53	22
Finland	42	23	50	18
France	–	–	46	17
Germany	–	–	31	9
Greece	–	–	15	6
Hungary	14	13	12	5
Ireland	10	9	41	17
Italy	14	8	26	8
Luxembourg	–	–	36	14
Netherlands	24	21	41	15
New Zealand	34	23	–	–
Norway	45	35	–	16
Poland	11	8	16	4
Portugal	–	–	17	7
Spain	–	–	25	11
Sweden	–	–	61	18
Switzerland	14	9	–	–
United Kingdom	44	22	49	13
United States	33	18	–	–

Source: OECD (2003).

Notes: Data in columns 1 and 2 are from the International Adult Literacy Survey (IALS) and refer to job-related education and training that employers provided (or partially paid for) to their workers aged 26 to 65 in 1994 for Canada, Ireland, the Netherlands, Poland, Switzerland (German- and French-speaking regions), and the United States, in 1996 for Australia, Belgium (Flanders only), New Zealand, and the United Kingdom, and in 1998 for the Czech Republic, Denmark, Finland, Hungary, Italy, Norway, and the Italian-speaking regions of Switzerland; data in columns 3 and 4 are from the European Continuing Vocational Training Survey (ECVTS) for 1999 and refer to firms with at least 10 employees; participation rate is ratio of employed persons in training to total employment (in percent); annual volume is hours per employed person.

The basic assumptions in the human capital model of education are threefold:

1. More education leads to higher productivity.
2. Higher productivity leads to a higher wage.
3. Individuals choose their level of education based on financial considerations.

The optimal level of schooling depends on the expected benefits and the costs of schooling. The expected benefits are determined by the higher wage and higher

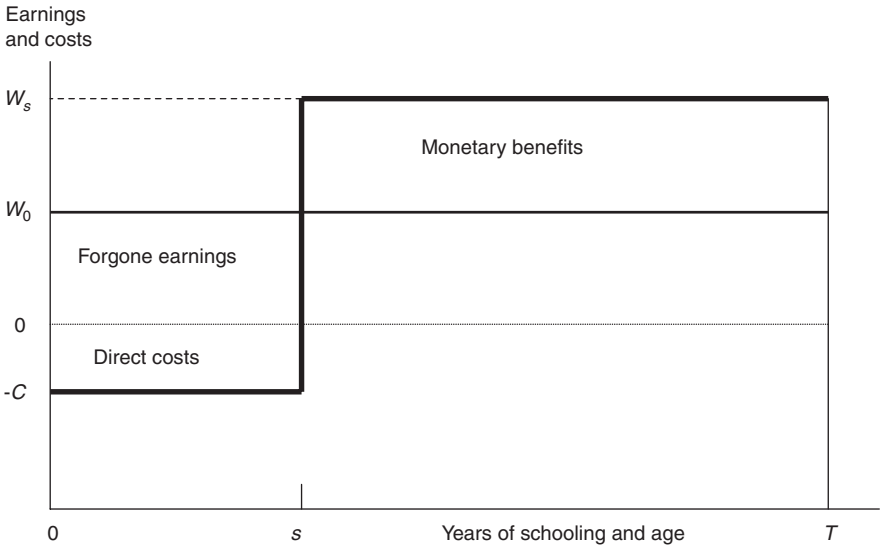


Figure 8.1 Benefits and Costs of Education

employment probability associated with a higher educational attainment and also by the expected length of time over which the returns to education materialize (the period between the time the education is finished and the time an individual retires).

Figure 8.1 represents the choice of the individual. The initial situation, which may represent high-school or primary education, is normalized to zero. The individual has to compare two options. The first option is to start working immediately and have annual earnings w_0 , which we consider to be constant for the remaining labor market years until retirement at age T .⁷ The second option is to attend s years of schooling during which direct educational costs have to be paid, and after that the wage will be $w_s > w_0$. The total costs of s years of schooling are the direct costs and the forgone earnings from 0 to s ; the monetary benefits are the higher earnings over the period s to T . The individual will do the extra years of schooling if the present value of doing so is larger than the present value of not doing this (see technical annex 8.8.1 for details).

Attending a particular level of schooling is more attractive the longer the expected payout period. Therefore, typically young people go to school. The perceived costs of schooling are the sum of direct expenses related to education and the forgone earnings. After all, going to school is difficult to combine with a full-time job. By comparing the net present value of the future stream of earnings with and without attending a specific number of years of schooling, individuals

⁷ Note that we also ignore periods of unemployment.

may find out how large their rate of return to schooling is. If the market rate of return to investment (the *external* rate) is higher than the rate of return to schooling (the *internal* rate), the individual should reduce the number of years of schooling; if the market rate of return to investment is lower, the years of schooling should be expanded. The internal rate of return will depend on the ability of the individual and the costs of funding study. Thus different individuals will have different rates of return to schooling and will choose different levels of schooling accordingly.

The first of the three basic assumptions underlying the human capital model is that education increases productivity. This need not be the case. If employers find it difficult to establish the productivity of workers, educational achievements may be used as a signaling device even if education itself does not increase productivity. A particular level of education may be a signal of innate ability because abler workers will find it less expensive to reach a particular level of education. If employers use education to screen workers, it will still be worthwhile for high-ability individuals to attain a high educational level because the signal attached to it pays off. Thus the private returns to schooling are still there. However, the social rates of return to education are much lower. If education does not increase productivity, the only benefit from educational systems is that it helps employers find the employees they are looking for.

8.2.2 Training

After finishing formal schooling, individuals may increase their productivity through on-the-job training. On-the-job training is part of human capital theory (see Leuven 2005 for an overview). Here the issue is whether the worker or the firm pays for the training. Becker (1964) distinguishes between general training and firm-specific training. General training in a firm increases productivity, but it can also be used by other firms. Because workers are free to leave, firms can never recoup their investment in general training. Therefore, workers have to pay for this, for example, by accepting wages lower than their productivity during training. The productivity increase from firm-specific training cannot be used in the market, and firms will be willing to pay for it. If workers are not credit constrained, in a competitive labor market there is no market failure in training (Acemoglu and Pischke 1999).

A major problem with this theory is the empirical prediction that firms do not pay for general training. This is at odds with the empirical observation that few skills are firm specific and most skills may be industry specific. Furthermore, there is evidence that although many training programs are general, firms pay a major parts of the costs of these programs. Acemoglu (1997) argues that because of labor market imperfections, employers have some monopsony power (see chapter 2) through which they can reap some of the benefits of general training causing a

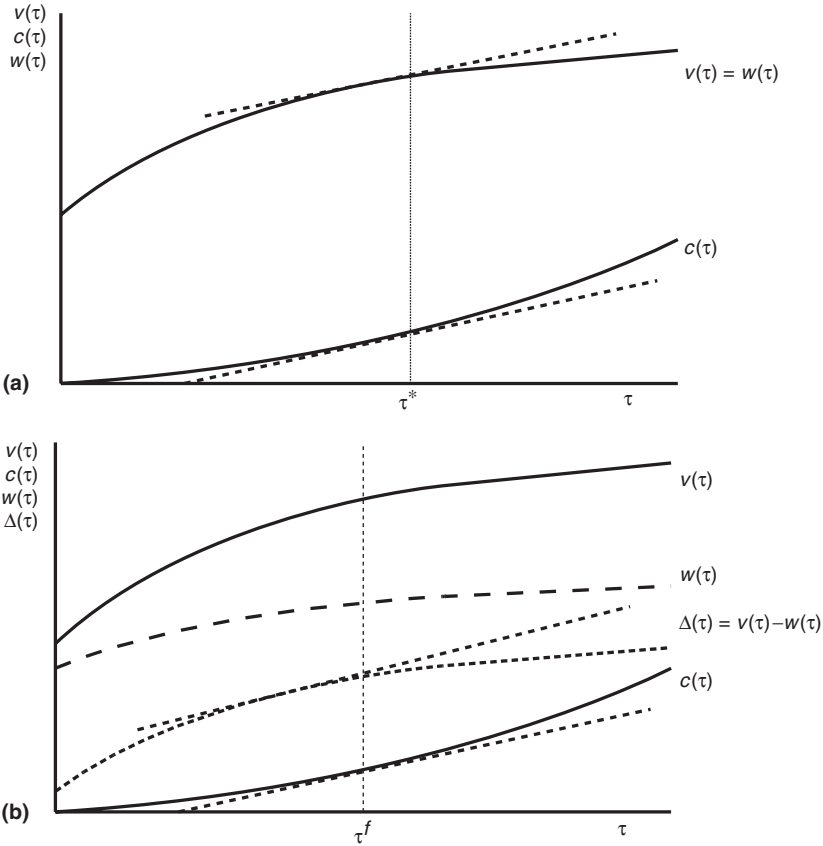


Figure 8.2 Training: Competitive Markets—Choice of Workers (a, Upper Graph); Noncompetitive Markets—Choice of the Firm (b, Lower Graph)

compressed wage structure. The line of reasoning is as follows (Acemoglu and Pischke 1998; see also technical annex 8.8.2). Irrespective of the market structure, the amount of general training τ increases the productivity and thus the output of the worker, $v(\tau)$, which is therefore upward sloping (see the upper graph in figure 8.2). But there are diminishing returns to training. To reward the worker for the increased productivity, wages also increase with the amount of general training, so the wage, $w(\tau)$, is also upward sloping with diminishing returns. Training costs will also increase with the amount of training, but increasingly so.

In competitive markets firms compete for the labor services of workers, and therefore $w(\tau) = v(\tau)$; hence the firm does not benefit from the general training. This being the case, it is obvious that the firm will not pay for the training, and the worker has to bear the costs. The choice of the worker is depicted in the upper graph of figure 8.2. An optimizing worker will choose the level of training

such that the difference between production and costs is at its maximum, that is, where marginal costs equal marginal benefits, which occurs when $\tau = \tau^*$. In this situation the worker benefits from the training, and therefore it is optimal for the worker to pay for it. However, the worker's choice may be constrained. A requirement for the worker to make the investment is that the resources to invest in training are accessible. This may not be the case. In the presence of credit market restrictions, the worker may not be able to make the investment. Since the employer cannot make the investment either, there is a market failure. In this case the government may want to intervene and provide subsidies for general training.

Now let us consider the situation in a monopsonistic market. In such a market the firm is able to pay the worker below productivity, and there is a gap: $\Delta(\tau) = v(\tau) - w(\tau)$. We assume that the gap between the value of labor productivity and the wage is increasing with the level of training. This implies that the returns to training for a worker are less than those prevailing in a competitive market; the wage structure is compressed. This situation is depicted in the lower part of figure 8.2. Now the firm has an incentive to invest. The gap between productivity and wages increases with the level of training, as does the cost of training. An optimizing firm will choose the level of training that equals marginal costs or marginal benefits, which now occurs when $\tau = \tau^f$. Thus, in a noncompetitive market a firm is willing to pay for general training of its workers because it is profitable to do so. In a noncompetitive market workers are not paid for their full marginal product even though their skills are general; that is, general skills are rewarded as if they were specific skills.

As also discussed in chapter 2, a firm may have monopsony power for various reasons. It may be that there are transaction costs in the labor market due to matching and search frictions. This may make it difficult for workers to leave their existing job and find a new employer to take full benefit of the general training. It may also be that there is asymmetric information between the current employer of the worker and other firms. The current employer knows the exact content of the training, while the other firms have incomplete information. Therefore, they are not willing to pay the worker according to that worker's marginal productivity. This would imply that in the lower part of Figure 8.2, the upward sloping $w(\tau)$ only refers to wages at the current employer, whereas in the labor market $w(\tau)$ would be a horizontal line. All in all, the monopsony power of the firm makes investment in general training profitable.

8.3 Empirical Evidence

8.3.1 Returns to Schooling

Table 8.3 shows that there is a strong relationship between educational attainment and labor market status. Employment rates are substantially higher for workers

Table 8.3 Employment Rates and Relative Earnings by Educational Attainment, Men and Women Aged 25 to 64

	Employment rates						Relative earnings					
	Men			Women			Men			Women		
	1	2	3	1	2	3	1	2	3	1	2	3
Australia	74.1	86.3	89.5	51.0	67.2	77.8	84	100	142	84	100	146
Austria	64.0	80.4	84.8	45.3	67.0	79.1	–	–	–	–	–	–
Belgium	61.7	81.7	87.9	35.8	63.9	79.8	90	100	132	81	100	132
Canada	67.1	82.2	86.3	46.2	70.6	78.7	82	100	139	67	100	136
Czech Republic	51.2	83.3	92.1	38.4	65.8	79.0	–	–	–	–	–	–
Denmark	70.5	84.3	87.9	54.7	74.3	85.2	87	100	131	90	100	123
Finland	60.2	77.5	87.0	53.1	70.5	82.5	92	100	163	98	100	146
France	68.0	81.4	85.6	52.0	68.5	78.3	88	100	159	81	100	146
Germany	59.7	74.6	85.6	41.4	64.3	78.4	90	100	150	81	100	145
Greece	78.3	84.8	88.6	36.9	53.3	75.3	–	–	–	–	–	–
Hungary	43.3	77.5	87.8	32.6	63.6	78.7	83	100	274	78	100	208
Iceland	86.1	91.5	94.4	76.4	81.9	92.3	–	–	–	–	–	–
Ireland	73.7	89.0	91.3	37.7	64.1	81.1	82	100	135	64	100	161
Italy	70.8	83.5	87.9	32.6	63.7	77.1	74	100	162	78	100	147
Japan	79.4	88.9	92.8	53.4	59.8	64.3	–	–	–	–	–	–
Korea	80.0	85.6	89.6	57.9	54.0	57.5	73	100	127	75	100	176
Luxembourg	72.6	82.9	89.5	47.4	54.9	74.5	–	–	–	–	–	–
Mexico	92.6	91.2	91.1	40.3	55.9	71.3	–	–	–	–	–	–
Netherlands	75.1	84.2	88.4	46.1	70.9	82.2	84	100	143	72	100	155
New Zealand	76.9	89.7	90.3	54.9	73.3	79.0	78	100	135	80	100	132
Norway	70.0	82.7	90.8	54.1	74.6	87.9	86	100	139	83	100	141
Poland	45.2	68.3	85.5	31.0	54.0	79.8	–	–	–	–	–	–
Portugal	81.0	83.1	88.8	62.7	77.5	87.5	–	–	–	–	–	–
Slovak Republic	27.3	77.0	88.2	19.3	63.1	79.1	–	–	–	–	–	–
Spain	77.1	84.7	87.3	37.8	61.0	76.5	79	100	138	64	100	125
Sweden	72.7	83.2	85.9	59.4	78.1	85.0	90	100	144	91	100	132
Switzerland	78.1	88.3	93.3	57.9	73.0	82.5	77	100	138	76	100	156
Turkey	76.0	80.8	81.6	22.6	26.8	63.4	–	–	–	–	–	–
United Kingdom	60.1	84.9	89.4	46.9	73.9	85.6	73	100	151	70	100	180
United States	68.1	78.6	87.3	43.5	67.4	77.5	67	100	189	70	100	177

Sources: Educational attainment: OECD (2006a); relative earnings: OECD (2005a).

Notes: Educational attainment is the highest completed level of education, defined according to the International Standard Classification of Education (ISCED); 1 = less than upper secondary education; 2 = upper secondary education; 3 = tertiary education; data for 2004. Relative earnings are defined as mean annual earnings from employment of individuals with a certain level of educational attainment divided by the mean annual earnings from employment of individuals whose highest level of education is the upper secondary level; information years 2001–3.

with tertiary education. Across countries there is wide variation in employment rates among low-educated workers. Employment rates among low-educated men range from a low of 27.3 percent in the Slovak Republic to a high of 92.6 percent in Mexico, while for low-educated women the extremes are 19.3 percent in the Slovak Republic and 76.4 percent in Iceland. For highly educated individuals, the cross-country range in employment rates is substantially smaller. Among highly educated men the extremes are 81.6 percent in Turkey and 94.4 percent in Iceland. Among females the range is from 57.5 percent in Korea to 92.3 percent in Iceland.

Table 8.3 also shows that not only are employment rates higher among highly educated individuals, but also their earnings are substantially higher. Among men the largest range between the levels of education is in the United States. Here men with less than upper secondary education earn only 33 percent less than men with upper secondary education. In Hungary men with tertiary education earn 174 percent more than men with upper secondary education. In other countries these differences are smaller. In Finland men with lower education earn only 8 percent less than men with average education. In Denmark men with tertiary education earn only 31 percent more than men with average education. Also among women there are clear differences in earnings by educational attainment, but also for women the magnitude of these differences depends very much on the country they live in.

Traditionally, empirical evidence of the rate of return to educational and training investments is derived by performing wage regressions with the amount of schooling and training as explanatory variables. In the analysis potential selectivity—high-ability individuals have more education—has to be accounted for. This is done by exploiting natural experiments using compulsory school laws, distance to school, draft lotteries, and other indicators as determinants of schooling unrelated to a person's ability. Alternatively, studies based on twins have been used to account for differences in ability across individuals. Identical twins are assumed to have the same ability. Hence wage differences between two members of a twin pair can be fully attributed to differences in educational attainments.

Ashenfelter and Rouse (1998) study wage differences between U.S. twins. They find the ability bias to be very small; cross-sectional estimates that ignore ability differences are only marginally upward biased (see box 8.1 for details). This result is similar to the findings of Isacsson (1999), who compares individual characteristics of Swedish identical twins to years of schooling. In his comparison he uses two physiological dimensions, weight and height, and two psychological dimensions, degrees of introversion and emotionality. He finds that the within-pair returns to schooling are 4.6 percent which is almost identical to the average of 4.5 percent estimated using a sample of the Swedish population at large.

Apparently, also in Sweden the effect of selectivity in schooling choice is not very important.

Box 8.1 Returns to Schooling: Using Data on Identical Twins

To establish the rate of return to schooling, one has to account for potential selectivity in the years of schooling. It could be that high-ability individuals have more education than low-ability individuals. If that is the case, the schooling effects measured in simple wage regressions represent a combination of schooling and ability. To correct for this, some studies use information about identical twins. The assumption is that identical twins have the same ability. Differences in earnings between twins may be related to differences in schooling but not to differences in ability. Wage regressions on the basis of identical twins can be used to estimate the true returns to schooling. Orley Ashenfelter and Cecilia Rouse do this on the basis of a U.S. sample of 340 twins. If they ignore the fact that the individuals in the sample are twins, they estimate a rate of return of 10.2 percent. If they base their estimates on a comparison of the twins, they find that the average return to schooling is 8.8 percent per year attained, so the ability bias is only 1.4 percent. Apparently, cross-sectional estimates that ignore ability differences are only marginally upward biased. A potential problem with twin studies is that the fundamental assumption of identical twins having identical abilities may not be fulfilled. Ashenfelter and Rouse find no evidence of this. They conclude that schooling investments of genetically equivalent individuals are the same, apart from random deviations that are used to estimate the returns to schooling. A Stata data file with Ashenfelter and Rouse's dataset and a program (do file) are available at the website (<http://press.princeton.edu/titles/8771.html>).

Source: Ashenfelter and Rouse (1998).

Angrist (1990) uses the U.S. government lottery during the Vietnam War, which established draft priority among potential draftees. Men with low lottery numbers were the first to be drafted, whereas men with higher numbers completely escaped the draft. Men with low lottery numbers who did not want to go to Vietnam had going to college as an alternative. The two groups were determined by a lottery and therefore on average had the same ability. Therefore, comparing men with low lottery numbers and high education to men with high lottery numbers and lower education reveals the true rate of return to schooling. Angrist and Krueger (1991) use compulsory school laws that sometimes forced workers to remain in school until they reached some predetermined age. Individuals under different regimes will have different years of schooling but are observationally equivalent in terms of ability. Comparing these individuals will generate an unbiased estimate of the rate of return to schooling. Card (1999) uses the distance to school to correct for selectivity in the schooling decision. The basic assumption is that parents decide where to live with their family without taking into account the distance

to school. Differences in distance to school are therefore unrelated to the ability of the children. Nevertheless, children who live closer to school will be more inclined to stay in school. Thus comparing individuals with different distances to school generates an unbiased estimate of the rate of return to schooling.

8.3.2 On-the-Job Training

In line with the prediction from human capital theory, a worker's wage is positively related to past investments in training. However, employers almost always pay the costs of training, while most of the training received is general training (Loewenstein and Spletzer 1998). Other than empirical analysis of the returns to schooling, there are only a few studies that estimate rates of return to training. Frazis and Loewenstein (2005) derive estimates of the rate of return to formal training using U.S. data. They find that 60 hours of formal training, the median positive amount of training, increases wages by 34 percent. This implies that rates of return at median positive training of 60 hours are in the 150 to 175 percent range; after correcting for promotions, direct costs of training, and heterogeneity in wage growth and wage levels, rates of return are 30 to 40 percent. There appears to be much heterogeneity across jobs, with managers and professionals having higher rates of return than blue-collar workers. Booth and Bryan (2005) investigate the effects of training on wages in the United Kingdom and find that only employer-financed training has a positive effect on wages, mainly through the incidence of training with previous employers. Employer-financed training increases the wage nearly 10 percent.⁸ They also find that self-financed training has no effects on wages. Studies by Acemoglu and Pischke (1998, 1999) on training indicate that the existence of imperfect labor markets may explain why firms pay for general training (see also box 8.2 and technical annex 8.8.2).

Box 8.2 On-the-Job Training in Germany

In Germany firms voluntarily offer apprenticeships to workers entering the labor market. Firms that train apprentices have to follow a prescribed curriculum, and apprentices take a rigorous outside exam in their trade at the end of the apprenticeship. Firms are not allowed to train apprentices without permission of the industry or crafts chambers. The training is monitored by worker councils in the firm. All in all, firms have limited influence on the training content. Therefore, most of the skills learned during the apprenticeships constitute general training. Nevertheless, contrary to the standard prediction of human capital theory, much of the financial

(continued)

⁸ Note that the fact that training by previous employers matters indicates that this training must have been general training.

Box 8.2 (continued)

burden of the apprenticeships is borne by the firm. Daron Acemoglu and Steve Pischke explain this phenomenon through the existence of labor market imperfections: mobility of workers is restricted, and therefore employers earn rents on trained workers. Given that firms are able to obtain part of the marginal product of workers, they have an interest in increasing this by investing in the human capital of the workers. Because of the monopsony power of the firm, workers are not willing to pay for general training themselves because they realize that a large fraction of the returns on the training will be appropriated by the firm. Acemoglu and Pischke compare wages of about 5,000 German apprentices who stay in their firm with those who quit for an exogenous reason, that is, go into the army. Their model predicts that military quitters should have higher wages than stayers because the exogenous reason of their separation makes them free from the monopsony power of the firm. Since the military conducts a physical exam but no other tests to determine draft eligibility, there is no selection bias. Furthermore, military service does not increase human capital of the draftees to an extent that is valued in the civilian labor market. They compare the gross monthly wages of the two groups with those of voluntary quitters from their firm and find for stayers a wage increase of 1.2 percent and for (military) quitters 4.5 percent. The model of Acemoglu and Pischke allows for multiple equilibria. One equilibrium is where workers who quit receive low wages. Therefore, firms will have considerable monopsony power and support a high level of training. Another equilibrium is where workers who quit receive high wages. This limits the monopsony power of firms, and the investment in training will be small. Paradoxically, a better allocation of workers to jobs—the equilibrium with high quits—may be less efficient because the level of training is lower.

Source: Acemoglu and Pischke (1998).

8.4 Policy Issues

8.4.1 Should Governments Subsidize In-Company Training?

Governments may stimulate training within firms through direct subsidies or tax credits. This will make it cheaper for firms to provide training. The main question is whether such support is optimal from a welfare point of view. If many firms would provide training anyway, the deadweight loss is large and the social returns to the investment are low. The answer to the question will depend on the market power of firms. In labor markets with imperfect competition, firms may have sufficient market power to provide training themselves without governmental support. In a more competitive market, employers may be reluctant to invest in training because other firms in the same industry may reap some or all of the benefits of the investment. If the in-company training increases productivity, there are social returns to training. Thus, if the social returns are larger than the private returns, governments may step in and subsidize training.

8.4.2 Should There Be a Compulsory Schooling Age?

In all OECD countries there is a compulsory schooling age that forces individuals to stay in school up to some minimum age. The question is whether such a compulsory schooling age is welfare improving. After all, if individuals drop out of school early, they may have good reasons to do so, that is, because they face negative returns to additional schooling. Nevertheless, a compulsory schooling age may be welfare improving if individuals are shortsighted and do not take into account all the future revenues of additional schooling. They may have a discount rate that is too large, for example, because they ignore the negative correlation between educational attainment and lifetime unemployment rates. A compulsory schooling age may prevent individuals from dropping out. It may also be the case that students face restrictions on the funding of their study and therefore drop out. Here too it may be that the social returns to schooling are higher than the private returns. Having a more highly educated population has positive externalities. In that case having a compulsory schooling age makes sense.

8.5 Why Do Governments Provide Education and Training?

Education and training increase labor productivity. Having a well-educated and well-trained workforce is an important competitive asset. The main reasons for governments to intervene in education and training are the presence of market imperfections and the positive externalities related to a more highly educated workforce. Individuals who want to pursue their studies will not always find it easy to finance this. Capital market imperfections may make it impossible or expensive to borrow money. This may lead to a suboptimal investment in human capital. There are private returns to schooling because individuals who invest in their education increase their productivity and will receive higher lifetime earnings. Because productivity goes up, education also has a social return. Through the investment in schooling, national income goes up. The social rate of return to schooling may be higher than the private return to schooling, for example, if individuals are credit constrained. This may be an argument for governments to subsidize formal educational systems. Without subsidies there would be an underinvestment in schooling. Positive externalities provide an additional reason for governments to provide education and training. Human capital, with education as its main determinant, is one of the main sources of economic growth. Firms may be reluctant to invest in training if this training has productive effects outside the firm as well. A firm will only invest in training if it has sufficient market power to reap at least part of the benefits. Governments may want to subsidize training if the social returns to training are higher than the private returns. Thus underinvestment in human capital is prevented.

8.6 Suggestions for Further Reading

The OECD study *Education at a Glance* (2005a) provides an interesting overview of cross-country differences in educational systems. Orley Ashenfelter and Cecilia Rouse (1998) discuss methodological issues of establishing the rates of return to schooling. Daron Acemoglu and Steve Pischke (1998) discuss job training and who should pay what in which situation. David Card (1999) surveys literature on the causal relationship between education and earnings. Giorgio Brunello, Pietro Garibaldi, and Etienne Wasmer (2007) provide a theoretical and empirical analysis of education and training in Europe.

8.7 Review Questions

1. Why do firms pay for general training even though trained workers are valuable for other firms as well?
2. Why do governments provide scholarships to students even though education is an investment from which students themselves will benefit later in life?
3. Why should there be a compulsory schooling age?
4. Why is it difficult to measure returns to schooling?
5. How are compulsory school laws or distance from schools used to derive unbiased estimates of returns to schooling?
6. Why should not all students try to achieve an academic degree?
7. Does it matter for the schooling decisions of the individual to what degree schooling is a signal of innate productivity?
8. Should the state subsidize on-the-job training?
9. Why are employment rates positively correlated with educational level?

8.8 Technical Annex: Schooling and Training

8.8.1 Optimal Years of Schooling

In figure 8.1 an individual chooses between an initial number of years of schooling (which might be primary schooling but is normalized to zero) and a particular number of years of schooling s . Alternatively, we may consider the decision whether an individual who has s years of schooling should opt for another year of schooling. If the individual does not attend this additional year, the annual earnings over the remaining labor market years until age T will be w_s . If the individual attends $s + 1$ years of schooling, the annual earnings from age $s + 1$ to age T are equal to w_{s+1} . The annual costs involved at school are equal to c_s . The individual makes a decision on the basis of a comparison of present values of both

flows of earnings and costs. If i is the market interest rate, $\delta = \frac{1}{1+i}$ is the discount rate used by the individual. The net present value with s years of schooling is

$$NPV_s = \sum_{a=0}^T \delta^a w_s = w_s + \sum_{a=1}^T \delta^a w_s. \quad (8.1)$$

The present value with $s + 1$ years of schooling is

$$NPV_{s+1} = -c_s + \sum_{a=1}^T \delta^a w_{s+1}. \quad (8.2)$$

The individual will attend another year of schooling if the present value of doing so is larger than the present value of abstaining: $NPV_{s+1} > NPV_s$. After some rewriting, this leads to

$$\sum_{a=1}^T \delta^a (w_{s+1} - w_s) > w_s + c_s. \quad (8.3)$$

If the costs of schooling are small, $c_s \approx 0$, and again after some rewriting it follows that:

$$w_{s+1} - w_s > w_s \frac{1 - \delta^9}{\delta}, \quad (8.4)$$

from which it follows that $w_{s+1} > \frac{w_s}{\delta}$. So

$$\ln(w_{s+1}) > \ln(w_s) - \ln(\delta) = \ln(w_s) + \ln(1+i) \approx \ln(w_s) + i, \quad (8.5)$$

and therefore

$$\ln(w_{s+1}) - \ln(w_s) > i. \quad (8.6)$$

Condition 8.6 is easy to interpret since the left-hand side represents the rate of return to an additional year of schooling: $\ln(w_{s+1}) - \ln(w_s) = r$. In other words, the individual will do the additional year of schooling if $r > i$, if the rate of return to schooling at educational level s is larger than the market interest rate. The individual will invest in schooling if it is more profitable than an investment in the capital market. It may be that the rate of return to schooling decreases with the level of schooling. Since the market interest rate is independent of the level of schooling, the individual will have to make a choice concerning the optimal number of years of schooling, s^* . Optimizing behavior implies that the individual will choose s^* such that $i = r$. The individual will keep investing until the marginal rate of return to schooling equals the market interest rate.

⁹ Using the approximation $\sum_{a=1}^T \delta^a \approx \frac{\delta}{1-\delta}$ when T is large.

8.8.2 Who Pays for General Training?

The answer to the question of who pays for general training depends on the competitiveness of the labor market. Assume that the productivity of a worker depends on the amount of general training, τ :

$$v(\tau) = \bar{v} + v_0(\tau), \quad (8.7)$$

where $\bar{v}$ is the productivity without general training and $\frac{\partial v}{\partial \tau} > 0$, $\frac{\partial^2 v}{\partial \tau^2} < 0$. Also the wage is assumed to depend on the amount of general training:

$$w(\tau) = \bar{w} + w_0(\tau), \quad (8.8)$$

where $\bar{w}$ is the wage without general training and $\frac{\partial w}{\partial \tau} > 0$, $\frac{\partial^2 w}{\partial \tau^2} < 0$. Thus a firm's profits Δ from an individual worker depend on the amount of general training:

$$\Delta(\tau) = v(\tau) - w(\tau) = \bar{v} - \bar{w} + v_0(\tau) - w_0(\tau), \quad (8.9)$$

and the effect of the amount of training on the profits of the firm is given by

$$\frac{\partial \Delta}{\partial \tau} = \Delta'(\tau) = v'(\tau) - w'(\tau). \quad (8.10)$$

Now in a situation where there is no general training ($\tau = 0$), two possibilities can be distinguished:

1. The labor market is competitive, so $\Delta'(0) = 0$ and the firm does not benefit from general training. However, the worker has an incentive to invest in training because the wage increases with the amount of general training. The worker chooses the optimal amount of investment τ^* by maximizing the difference between the wage and the costs of general training, $c(\tau)$, with $\frac{\partial c}{\partial \tau} > 0$ and $\frac{\partial^2 c}{\partial \tau^2} > 0$. This leads to $w'(\tau^*) = c'(\tau^*)$, a situation that is depicted in the upper part of figure 8.2.¹⁰
2. The labor market is noncompetitive and there is wage compression, so $\Delta'(0) > 0$: the productivity increases faster than the wage. Now the firm has an incentive to invest in general training because this increases the firm's profits. The firm chooses the optimal amount of investment τ^f by maximizing the difference between profits and costs of general training, which leads to $\Delta'(\tau^f) = c'(\tau^f)$, a situation that is depicted in the lower part of figure 8.2.

¹⁰ In figure 8.2 it is also the case that $\bar{f} = \bar{w}$, but of course this is not necessary for the general result to hold.

Migration Policies

International migration is the great absentee in the era of globalization. While the barriers to international trade and capital mobility have already been largely removed, cross-border worker flows are tightly restricted. Until a few decades ago many European countries were mass out-migration countries. According to historians, about 60 million Europeans moved away from the Old Continent in the period 1820–1940. Two-thirds of them went to the United States (one can find their names on the website <http://www.ellislandrecords.org>). Currently Europe is attracting more migrants in proportion to its population than the United States. Countries like Italy that were sending 1 to 2 million people to the United States every year at the beginning of the twentieth century are now receiving annual inflows of migrants of the order of 350,000. In Europe migration policies are getting stricter and stricter. At the same time, poor enforcement of these restrictions is giving rise to very large inflows of illegal migrants. According to some estimates, up to 40 percent of migration to Europe is illegal.

Migration policies introduce a complex set of restrictions on the movement of persons across jurisdictions. Most typically, migration restrictions take the form of *quotas* that establish a maximum number of work and residence permits to be issued to foreigners in a given year. Quotas are often allocated on a first-come, first-served basis, but an increasing number of countries are adopting a *points system*, whereby each application is attributed a score based on explicit criteria that typically reward educational attainment, experience, and language abilities. Bonus points can also be given for employment in occupations and regions where there is a shortage of workers. Migration is also deterred by increasing the administrative burdens placed on migrants and their employers. The number of bureaucracies involved in *admission procedures* is therefore another indicator of the stance of migration policies. Another important feature of migration legislation relates to the *length of stay*, that is, the maximum duration of the residence or work permit. Policies related to the assimilation of migrants are also relevant in defining the more or less restrictive stance of migration policies. A good indicator in this context is the minimum number of years required to apply for *citizenship*. Finally, there are ad hoc rules for *asylum seekers*. Although asylum policies are largely inspired by noneconomic considerations, in practice most asylum seekers respond

to economic incentives, just like other migrants. Hence it is important to include asylum policies when analyzing the more or less restrictive stance of any given country toward international migration.

Economic theory suggests that migration increases efficiency by arbitraging away cross-country (and even cross-regional when natives are immobile) differences in productivity and unemployment. The scope for this efficiency-enhancing arbitrage is large: in international markets for commodities and financial assets, price differentials rarely exceed a ratio of 2/1, while wages of similarly qualified individuals in advanced and low-income countries differ by a factor of 10 or more. This suggests that gains from liberalizing labor movements across countries are enormous and much larger than the likely benefits from further liberalization in the traditional areas of goods and capital. Why then is migration so severely restricted? The answer is that migration policies are essentially redistributive tools, aiming at reducing negative effects of migration on wages and unemployment among natives. Moreover, migration tends to occur in waves. By imposing some gradualism on migration flows, migration restrictions mitigate supply shocks that may negatively affect incomes or jobs of some specific groups.

In addition to aiming at controlling the magnitude of migration flows, migration policies can also affect their composition, for example, by encouraging more or less skilled migration. For this reason, in most countries migration policy is a highly controversial topic. There are, on the one hand, those who are keen to host more migrants, who need them desperately to fill vacant positions or to find someone who can pay for their pensions, and, on the other hand, those who fear that migrants will ultimately steal their jobs and increase crime rates and abuse of social transfers paid out of their pockets. The same individual may at the same time support and oppose migration, and vigorously do so. Employers wish to have migrants from 7 A.M. to 5 P.M., rigorously just during working hours, and then, when it is time to shift to their private life, prefer not to see them any longer. As in the famous aphorism on migration to Switzerland by Max Frisch, “We wished arms, and we received persons.”

Migration policies have relevant interactions with other institutions discussed in this volume. In particular, they are affected by minimum wage legislation (chapter 2), unions (chapter 3), family policies (chapter 7) and unemployment benefits (chapter 11) because wage compression and income floors affect the skill composition, and hence the fiscal and distributional effects, of migration.

9.1 Measures and Cross-Country Comparisons

It may be useful to put numbers in a historical context. Contrary to common wisdom, migration flows have not been increasing in the era of globalization. The United States is currently experiencing much lower inflows of migrants, in proportion to the native population, than at the beginning of the twentieth century (figure 9.1).

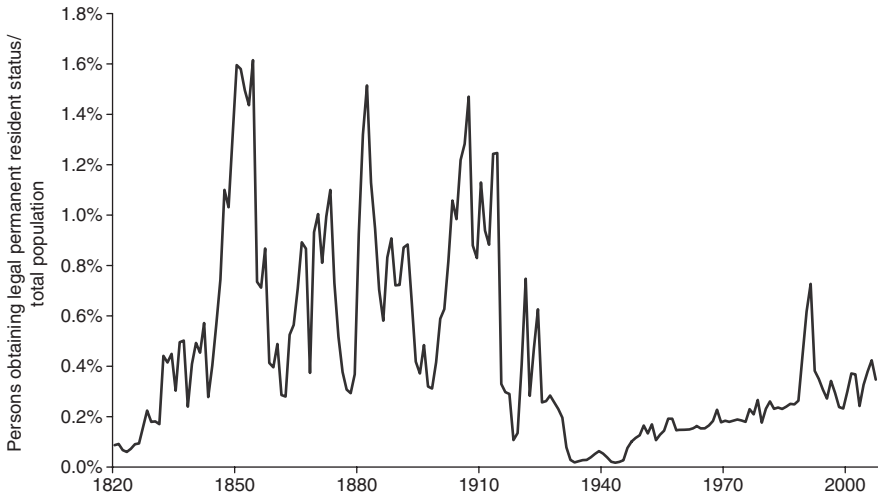


Figure 9.1 Gross Inflows of Legal Migrants to the United States as a Fraction of the U.S. Population, 1820–2007

Sources: U.S. Department of Homeland Security (for immigration data); and U.S. Census Bureau (for population statistics).

Europe experienced larger migration flows after the massive political disruptions of World War II and decolonization than in the past two decades. The number of people displaced by World War II was estimated at around 20 million. In the United Kingdom there was increasing immigration from the New Commonwealth after the war, beginning with the Caribbean countries and, later, the Indian subcontinent. In France more than 1 million Algerians of French origin were repatriated after independence. Similar processes occurred in Belgium, the Netherlands, and, later, in Portugal after the dissolution of their colonies.

Migration was not only tolerated but even encouraged in Europe in the 1950s. After full employment had been achieved, labor shortages induced some countries to open their labor markets or even to recruit foreign labor. The main source countries were in southern Europe (Italy, Greece, Portugal, Spain, Turkey, Yugoslavia) and North Africa (Morocco, Tunisia), while the main countries of destination were France, Germany, the United Kingdom, Switzerland, Belgium, and the Netherlands. The total number of foreign-born persons increased in Europe from around 4 million to 10 million from 1950 to 1970, while about 5 million workers migrated from southern Europe to these countries (Zimmermann 2005).

This liberal stance toward migration was abandoned at the beginning of the 1970s, when European unemployment was increasing. Since then, and even more in the last 15 years, migration policies have been getting stricter and stricter, not only in the North, but also in southern Europe.

Table 9.1 Strictness of Migration Policies in the Countries of the EU with the Largest Immigration

	Existence of quota System	Admission requirements	Years to obtain permanent residence	Residence requirements	Length of first stay	Overall index
		(1)	(2)	(3)	(4)	(5)
Austria	yes	5	5	3	2	2.8
Denmark	no	6	7	4	4	3.2
Finland	no	4	4	2	4	2.8
France	no	3	5	1	2	1.5
Germany	no	6	5	4	2	2.6
Greece	no	4	10	3	2	2.7
Ireland	no	6	10	3	2	2.9
Italy	yes	6	6	3	2	3.1
Netherlands	no	5	5	2	4	3.0
Portugal	yes	6	5/8	3	2	3.1
Spain	yes	6	5	2	2	3.2
UK	no	3	10/14	1	2	2.9

(1) The number of certificates and procedures needed to be admitted as a foreigner, whatever the motivations may be.

(2) The number of years required to obtain a permanent residence permit.

(3) The number of certifications or procedures required to legally reside in the territory. This differs from the requirements for entering the country because holding a valid document is typically not sufficient.

(4) The number of years of stay required to obtain a first residence permit.

(5) For details as to how the overall index is computed, see www.frdb.org.

Table 9.1 provides some information on migration policies in the countries of the EU with the largest immigration. As shown by the table, regulations vary considerably from country to country. The countries that are requiring a large number of admission requirements are generally also those that prolong the period to get permanent residence. A few countries adopt an explicit quota system. These are mainly the southern European countries that are facing immigration from North Africa.

Figure 9.2 tracks the evolution over time of a summary indicator of the stance of migration policies. The index (see www.frdb.org for details) is developed by normalizing countries' scores in each migration policy area in the 0–6 range and then taking the simple average of these subindicators. As shown by figure 9.2, most countries are above the bisecting line through the origin, pointing to a tightening of migration restrictions over time.

9.2 Theory

In competitive labor markets migration should negatively affect wages of natives. This effect comes from the presence of a downward-sloping demand curve. When labor supply is not rigid, migration also reduces employment among natives. The magnitude of these effects and the extent to which migration is accommodated

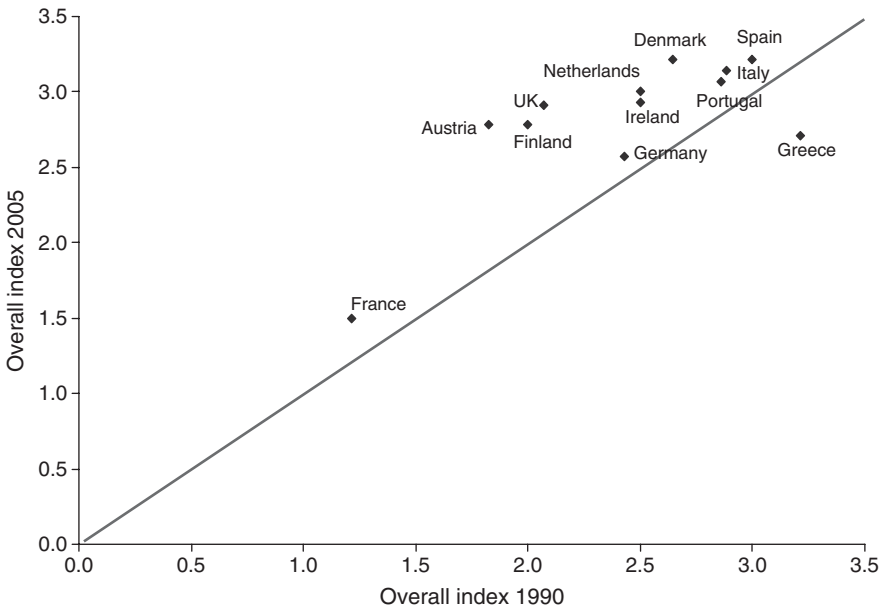


Figure 9.2 Evolution of the Index of Strictness of Migration Policies in the EU-15

Note: The index is the simple average of subindicators obtained along the policy dimensions characterized in table 9.1. See www.frd.org.

via reductions in wages or employment depends on the elasticity of labor supply in the relevant range. Labor market institutions that interfere with employment and wage adjustment are also important in this context. We begin, as usual, by characterizing the case of a competitive labor market where there are no other institutions (in addition to migration restrictions) interfering with labor market adjustment and then move to more realistic institutional configurations.

9.2.1 A Competitive Labor Market

Figure 9.3 depicts the effects of migration on a competitive labor market when labor supply is vertical (left-hand panel) or positively sloped (right-hand panel). For the time being, we assume that natives and migrants are perfect substitutes; that is, they are exchangeable. Later (section 9.2.5) we will consider skill differences and hence also complementarities between migrant and native labor that potentially shift the labor demand schedule. We will also focus on the short-run effects of migration, neglecting changes in the capital stock associated with the decline in wages in the receiving country.

Migration in this context is a shock to labor supply that shifts the aggregate supply curve to the right from L_0 , where labor supply involves only native workers, to

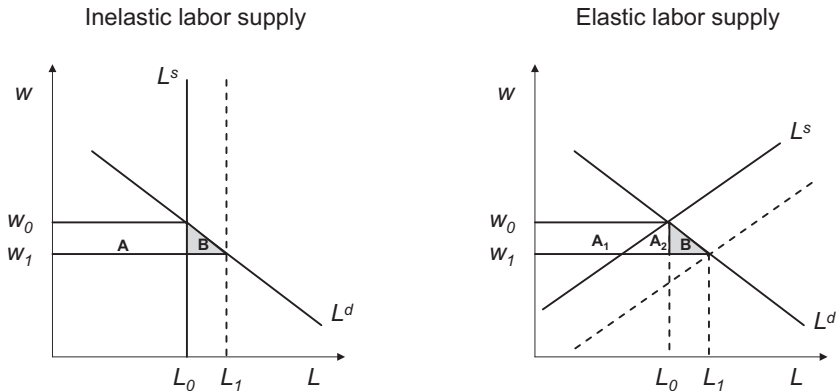


Figure 9.3 A Competitive Labor Market

L_1 , where there are both natives and migrants. When labor supply is rigid, there is no reduction of employment among natives, and the adjustment is accommodated entirely via reductions in wages. When instead labor supply is positively sloped, some reduction of employment among natives is also involved. Importantly, the reduction of employment among natives comes entirely from the withdrawal from the labor market of natives who were willing to work only at the wage levels prevailing before the migration-supply shock. No unemployment is involved in the adjustment.

Notice that in both cases native workers suffer a loss in welfare (in the left-hand side panel, their loss is given by the area A ; in the right-hand side, by A_1), while employers' gains ($A + B$ in the first case, and $A_1 + A_2 + B$ in the second) exceed labor losses. Thus there is an aggregate welfare increase in both cases. There is also an increase in national income accruing to natives, that is, an *immigration surplus* (Borjas, 2003) given in both cases by the triangle B . As shown by figure 9.3, this immigration surplus can be estimated as the product $\frac{1}{2}(w_0 - w_1)(L_1 - L_0)$, which suggests that it is increasing with the changes induced by migration to native wages and with the size of migration flows. At the same time, there is a conflict of interest between native workers and native employers over migration policy. In so far as migrants do not carry out any capital with them, labor productivity and income per capita decline in the recipient country.

9.2.2 An Economy with Wage Rigidities

Migration most frequently occurs to countries that impose in one way or another some form of wage rigidity. The case of an economy where downward wage adjustment is constrained by wage-setting institutions is characterized in figure 9.4. The panel on the left-hand side depicts an economy with a minimum wage set at the market-clearing level. Thus initially there is no unemployment. The supply shock

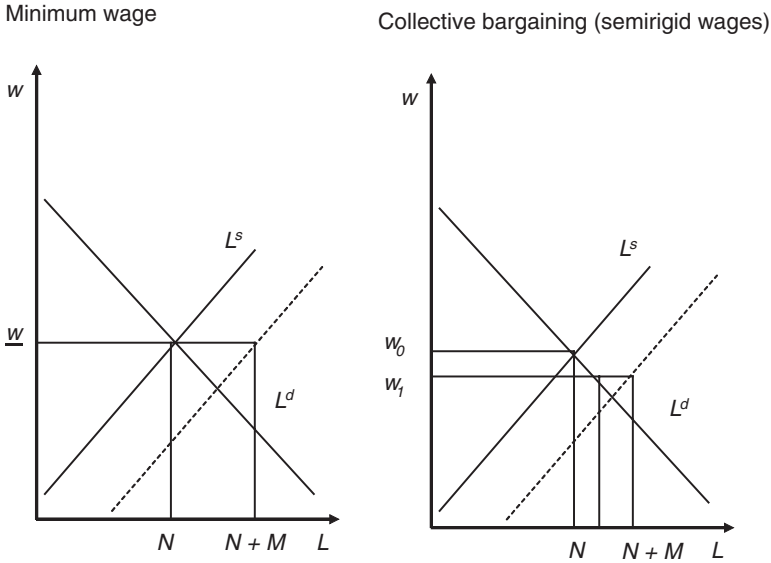


Figure 9.4 An Economy with Wage Rigidities

associated with migration in this case needs to be accommodated entirely by job losses. The extent of native worker losses will depend on the market penetration of migrant workers. With a binding minimum wage, there is indeed a fixed number of jobs in the economy. The more migrants who come in and find a job, the higher the unemployment among natives. If there are no externalities associated with unemployment (e.g., fiscal externalities related to the payment of unemployment benefits), then employers realize the same surplus with and without migration, but native workers lose out. As a result, total incomes of natives in the immigration country decline, and unemployment appears among native workers.

A more realistic scenario of partial wage rigidity (wages adjust, but not enough to allow for market clearing after migration takes place) is depicted in the right-hand panel of figure 9.4. Here there is no longer a one-to-one relationship between migration and unemployment, and native employers do gain something from the inflow of migrants. Thus it is possible that in this case the immigration country realizes a net gain from immigration because the accrued surplus of employers hiring the migrants exceeds the welfare loss of native workers earning lower wages. Yet the net gain will be lower than in the competitive labor market case.

9.2.3 Wage Rigidities and Unemployment Benefits

The preceding discussion suggests that under realistic institutional configurations (with some rigidity in downward wage adjustment), immigration creates some unemployment. If natives and immigrants have access to unemployment benefits funded via taxes and/or payroll social security contributions, then there will be a

fiscal cost associated with migration. In this context the incomes of natives in the receiving country are affected by migration via three main channels:

1. Changes in wages
2. Changes in employment opportunities
3. Changes in the tax rate

The first two effects arise also in the absence of unemployment benefits, while the third is related to the presence of a welfare state and can be defined as a *fiscal externality* of migration. An additional complication induced by the presence of unemployment benefits is that they may attract more migrants with personal characteristics that make them likely to draw unemployment benefits. In other words, the presence of unemployment benefits may alter not only the size but also the composition of migration. This effect can be better understood by considering the determinants of migration decisions.

9.2.4 What Drives Migration Decisions?

Migration is not a random process. It is a rational choice that involves two decisions. The first decision relates to *whether* to migrate. The second decision is on *where* to relocate. The two decisions are clearly interdependent, but they can conveniently be characterized as sequential choices.

Let us consider first the decision on whether to move to another country. This choice is based on a cost-benefit analysis that weighs pros and cons of migration. Typically the costs of migration are frontloaded; they are all paid at the time of changing residence. These migration costs are very large because they include not only transportation and relocation costs but also the huge psychological costs related to abandoning the network of relatives. The stream of benefits associated with migration comes from the earning differentials between the destination country and the country of origin. As shown in technical annex 9.8, migration occurs if

$$\frac{(w_F - w_H)}{w_H} > \frac{C_0 (1 - \delta)}{w_H T^e}, \quad (9.1)$$

where $(w_F - w_H)$ denotes the wage differential between the potential destination (the foreign country) and the home country, C_0 is the frontloaded migration costs, δ is the subjective discount factor, and T^e is the expected length of the working life. Condition (9.1) suggests that the lower the costs of mobility, the younger the individual (the longer the time horizon in which the initial investment can provide returns in terms of earning differentials), and the larger the difference in wages between destination and origin countries, the higher the probability that the individual migrates.

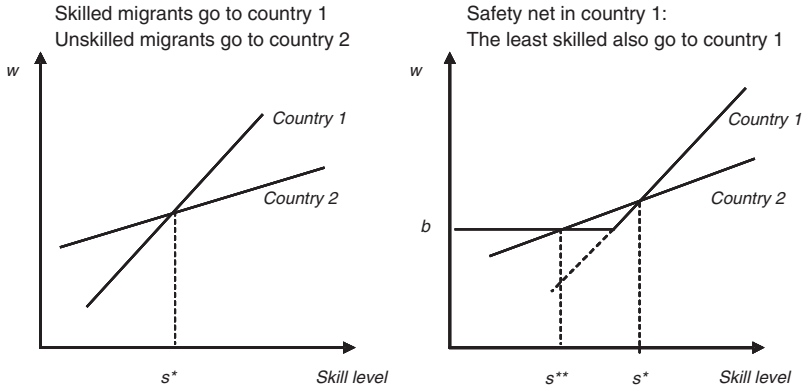


Figure 9.5 The Choice of Where to Migrate

The second decision relates to where to relocate. If we suppose for simplicity that the initial mobility costs are independent of the final destination (e.g., the individual is choosing among different states of the United States or within the EU), the choice of where to move will depend entirely on the potential earnings in the different destinations. Suppose that the migrant can choose between two alternative destinations, say, country 1 or 2, and that the only relevant dimension affecting the potential earnings of migrants is skill, as in the Roy (1951) model. The choice faced by the migrant is illustrated in figure 9.5, which displays on the horizontal axis the skill level of a given individual and on the vertical axis that individual's earnings, net of migration costs. As shown by the panel on the left-hand side, the most skilled workers (those with skills higher than s^*) will go to country 1, where their skills are better rewarded, while the least skilled will go to country 2 because they are paid more there.

Suppose now that country 1 introduces a welfare system in the form of a floor for incomes (e.g., a guaranteed minimum income scheme), preventing natives and migrants from falling below a given poverty threshold, say, b .¹ Now also some of the unskilled (those with skills lower than s^{**}) will migrate to country 1. Therefore, the presence of a welfare state affects the skill composition of migration and may induce some workers not to go where their skills are most productive.

9.2.5 Effects on Income Distribution of Skill-Biased Migration

As just discussed, the introduction of a safety net in a country may alter the skill composition of migration, increasing the proportion of unskilled migrants moving

¹ Notice that a minimum wage would not necessarily prevent income of migrants from falling below b , as it would create unemployment among migrants and natives.

there. Changes in the skill composition of migration may not only have important fiscal effects but can also affect the income distribution in the receiving country.

To illustrate this, we assume that there are just two types of labor, skilled and unskilled labor, and that migrants and native workers are perfectly substitutable only within each skill group. Suppose further that only one good is produced, combining both skilled and unskilled labor in a fixed proportion. This means that skilled and unskilled labor are complements.

If migration perfectly replicates the skill composition of the native population (migrants have exactly the same proportion of skilled and unskilled workers as the native population), then we should expect no effects of migration on relative wages and hence on income distribution in the recipient country. Suppose instead that all immigrants are unskilled workers, whereas in the destination country there are some skilled workers. The effects on wages of native unskilled workers will be just as described in sections 9.2.1 through 9.2.3: unskilled wages will decline, and, depending on the type and extent of wage rigidities, unemployment among unskilled workers may increase. At the same time, because of the higher number of unskilled workers in the economy, labor demand for skilled workers will shift out, and hence wages of skilled natives will increase. Thus wage differentials between skilled and unskilled workers in the receiving country will increase as a result of immigration of unskilled workers. The opposite clearly happens when immigration involves only skilled labor: in this case earnings inequality in the receiving country is reduced by migration.

This result is fairly robust. Relative wages will change in the direction outlined here even when some (imperfect) substitutability is allowed between skilled and unskilled labor (Dustmann, Fabbri, and Preston 2005). If we allow for two goods, both internationally traded and using skilled and unskilled labor in different proportions, cross-skill substitutability reduces the effects of skill-biased migration on income distribution in the long run as the country increasing its endowment of, say, unskilled workers gradually expands its production in the unskilled-labor-intensive industry (Leamer and Levinsohn 1995). But the key effect is still there, at least in the short run: unskilled migration increases income inequality, while skilled migration reduces income inequality in the recipient country. The type of labor that becomes relatively scarcer gains from migration, while the other loses.

9.3 Empirical Evidence

Most studies estimating the effects of migration on employment and wages of natives are based on U.S. data. This literature is reviewed in Borjas (2003). The empirical literature on European migration (reviewed in Boeri, Hanson, and McCormick 2002) is more recent because there is still a paucity of data on migration, notably illegal migration, to the EU. More empirical work outside the United

Table 9.2 Unemployment Rates of Native-Born and Foreign-Born Workers

Unemployment rates, workers 15–64 years old (2004)						
	Men			Women		
	Native-born	Foreign-born	Ratio	Native-born	Foreign-born	Ratio
Australia	6.0	6.5	1.1	6.1	6.5	1.1
Austria	4.3	11.2	2.6	4.3	10.7	2.5
Belgium	5.6	15	2.7	7.5	15	2.0
Canada	5.5	6.6	1.2	4.9	6.8	1.4
Denmark	4.4	14.4	3.3	5.2	10.3	2.0
France	8.0	13.6	1.7	9.9	17.2	1.7
Germany	10.3	18.3	1.8	9.6	15.2	1.6
Greece	6.5	6.7	1.0	15.7	18.9	1.2
Ireland	4.9	6.5	1.3	3.6	5.0	1.4
Italy	6.4	6.1	1.0	10.1	13.1	1.3
Netherlands	3.6	10.3	2.9	4.3	10.6	2.5
Portugal	5.7	9.9	1.7	7.4	9.6	1.3
Spain	7.8	11.4	1.5	15.1	17.1	1.1
Sweden	6.2	13.9	2.2	5.2	12.2	2.3
UK	4.7	7.3	1.5	3.9	7.3	1.9
USA	6.9	5.8	0.8	5.5	6.8	1.2

Source: OECD (2006a).

States is warranted to better characterize the effects of migration under different institutional configurations. As shown by table 9.2, unemployment rates of foreign-born males in Europe are generally higher than among the native-born population, while in the United States it is the other way around. Unemployment among women is somewhat less concentrated among migrants, perhaps because of the greater role played by nonparticipation of women in the migrant population. This suggests that cross-atlantic comparisons can be quite revealing about the consequences of immigration on the labor market.

The typical framework used in these studies correlates wages or unemployment among natives with some measure of immigrant penetration in a given area and often by skill category. This area-analysis framework is justified by evidence that immigrants concentrate in some gateway regions where the labor market adjustment to migration can be better observed.

As discussed earlier, economic theory suggests that we should observe a negative correlation between immigration and native wages and/or employment. However, the empirical literature until recently found very modest, if any, effects of migration on the labor market of natives. Typical estimates are that a 10 percent increase in the fraction of immigrants in a region lowers native wages by less than 1 percent (often an amount not statistically different from 0). These

findings are well summarized in a number of influential surveys concluding that “the effect of immigration on the labor market outcomes of natives is small” (Friedberg and Hunt 1995) or that “the weight of the empirical evidence suggests that the impact of immigration on the wages of competing native workers is small” (Smith and Edmonston 1997). The empirical evidence from EU countries also points to a negligible impact of migration on wages and employment opportunities of natives.

These empirical results can be at least partly reconciled with economic theory when account is taken of three factors:

1. Self-selection of immigrants into high-wage regions
2. Changes in migration patterns of native workers
3. Changes in the regional output mix

The first factor stems from the fact that immigrants are not randomly distributed across labor markets. As discussed earlier, they decide to locate to a given area in light of earnings and employment opportunities offered in that area. Thus regions with higher wages or lower unemployment are likely to attract more migrants, generating a *positive* correlation between migration and native wages. Causality in this context goes the other way around: wages of natives affect the levels and composition of migration rather than vice versa. This suggests that appropriate econometric methods should be found to deal with this *endogeneity* (of migration) problem.

The second factor relates to the responses of native workers. As more migrants arrive in a given area, natives may respond by moving their labor or capital to other regions or cities, thereby reequilibrating the national economy. This effect is likely to be important in the United States, where interregional labor mobility is relatively large (Topel 1986; Blanchard and Katz 1992) and strongly responsive to economic incentives. In Europe, by contrast, interregional labor mobility appears to be much lower (Boeri, Hanson, and McCormick 2002), and, if anything, foreign migrants compensate for the lack of mobility of residents. Figure 9.6 documents this fact: in Italy mobility from high-unemployment to low-unemployment regions is almost entirely accommodated by net migration of foreign workers as opposed to native workers.

Finally, the third explanation for the modest effects of migration on wages of natives draws from the fact that open economies may adjust to migration by changing the composition of the output mix or production technologies (Lewis 2005) in such a way as to reap the benefits from the increased labor supply, rather than by adjusting wages or employment to an increasing supply of labor. This adjustment in product specialization takes time and hence may play some role in explaining the small effects of migration on wages and employment of natives in low-frequency data.

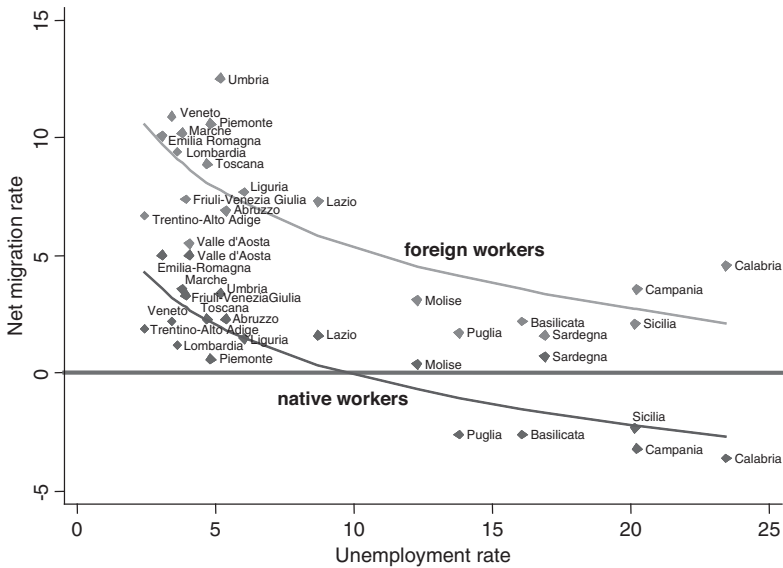


Figure 9.6 Net Internal and International Migration, Italy, 2003

For these three reasons, the effects of migration on labor market outcomes of natives can be better identified at the national level (Borjas 2003) and by correlating wages (or unemployment) of natives to the share of migrants by skill group. Results from the United States, Canada, and Mexico adopting this empirical framework (Borjas 2006) seem to indicate that migration has had a significant and negative effect on wages of natives, in line with the predictions of the competitive model of the labor market when capital is held constant. However, some of these national labor market studies focus on the total population rather than on the relative sizes of the different skill groups and use low-frequency data from the census (typically carried out at 10-year intervals). Card (2005), working on higher-frequency data on wages of school dropouts (people with less than a high-school diploma) in the United States, found that their wage gap with respect to other educational groups in the workforce had remained fairly constant in spite of substantial inflows of unskilled migrants.

Regrettably, only a few studies have used difference-in-differences techniques to isolate the effects of migration on wages and employment of natives. Among these few exceptions, Card (1990) analyzed Cuban migration to Florida following the opening of borders, Hunt (1992) studied migration to France after Algerian independence, and Carrington and de Lima (1996) investigated the effects of the inflows to Portugal of about 600,000 *retornados* after the loss of the African colonies of Angola and Mozambique. Card in particular documented that although the 1980 Mariel boatlift into the Miami-area labor market unexpectedly increased

the local labor supply by 7 percent in a four-month period, this shock had no discernible effect on native wages or unemployment in the area. These applications of the difference-in-differences techniques, however, raise a number of issues because it is not obvious how the treatment and control groups can be defined in the case of immigration. Native workers may indeed react to expected large immigration flows by moving somewhere else.

More convincing applications to migration decisions of the difference-in-differences techniques can be found when the focus is on the decision to migrate rather than on the effects of migration. For instance, Costa and Kahn (2000), using difference-in-differences techniques, documented the strong responsiveness of internal migration in the United States to economic incentives, looking at residential choices of power couples (couples in which both spouses are college graduates) (see box 9.1).

Box 9.1 Residential Choices of Power Couples in the United States

Internal migration, like international migration, is nonrandom. A study by Dora Costa and Matthew Kahn (2000) applying difference-in-differences analysis to residential location of couples with different educational attainments well documents this fact. Costa and Kahn investigated location choices of so-called power couples, that is, couples in which both partners have at least a college degree. It may be difficult in local areas for both members of the couple to find suitable jobs. Hence one of the spouses may be forced to commute or to take a job for which that spouse is overeducated. In order to avoid this type of choice, power couples may decide to migrate to large metropolitan areas where the labor market for top jobs is relatively large and offers opportunities for both spouses. Table 9.3 compares location choices of power couples in the 1970s and in the 1990s. In this 20-year period the employment rate of women rose dramatically in the United States, creating the thin labor market problem of power couples. The second difference used in the study is the location decision made by lower-power couples, that is, couples in which neither spouse is a college graduate. As indicated by the third row in the third column of the table, being a power couple increases the probability of living in a large metropolitan area by 8.5 percentage points, just as economic theory would predict.

Table 9.3 Difference-in-differences analysis of location choices of married couples

	1970	1990	Δ
Power couples	14.6	34.8	20.2
Nonpower couples	8.3	20.0	11.7
Difference	6.3	14.8	8.5

Source: Costa and Kahn (2000).

9.3.1 The Labor Market Performance of Migrants

A number of studies have analyzed the labor market performance of migrants, notably their convergence to wages of natives (Borjas 1999). This literature is relevant also to assess the contribution of migration to the national economy. Although most of this literature is based on cross-sections, it would be preferable to use longitudinal data that follow the same individuals over time in assessing convergence. The general finding, in any event, is that income convergence is largely determined by the human capital characteristics of migrants. While relatively high skill levels translate into a relatively favorable labor market outcome of migrants, the labor market performance of low-skilled migrants lags persistently behind that of natives. Unfortunately, this literature is mainly confined to the United States because European countries rarely have long series of longitudinal data on migrants.

The U.S. literature also provides support for the nonrandom allocation of migrants to different areas in the country of destination. As pointed out by Borjas (2001), migrants grease the wheels of the labor market: they migrate to the most productive areas, arbitraging away differences in labor productivity. Hunt (2006), in her study of east-west migration in the context of German unification (box 9.2), also documents a nonrandom allocation of migrants by age and skill, very much in line with the predictions of the Roy model reviewed earlier. In particular, she found that decisions of young people to migrate strongly react to wage differentials between the sending and the receiving country, while older workers are more responsive to unemployment differentials.

Box 9.2 Learning from East-West Migration after German Reunification

Drawing on regional-level and individual data on east-west and west-east migration after German reunification, Jennifer Hunt (2006) investigated the determinants of migration for different socioeconomic groups and the role played by labor market conditions in the region of origin and in the region of destination. Her findings shed some light on the reasons why the German government (supported by the unions) preferred to support wage convergence between East and West Germany notwithstanding the large productivity differentials between the two regions. This resulted in large unemployment in the East, but reduced east-west migration. Economists had opposed this strategy, claiming that it would have been ineffective in containing migration because unemployment is a strong push factor inducing many workers to leave the high-unemployment region.

(continued)

Box 9.2 (continued)

Hunt estimated separate regressions of the migration rate (the fraction of individuals in any given age group moving from East to West Germany) by age and showed that young workers (who contribute the most to migration flows because of their longer planning horizon) are much more affected in their migration decisions by wage differentials than by unemployment in the country of origin, perhaps because they can count on family support at home (see table 9.4). Older workers are instead induced to leave after a dismissal and when unemployment is high. Table 9.4 displays the estimated elasticity of migration with respect to wages and unemployment in the origin and in the destination. For instance, the number reported in the first row and column of the table indicates that a 1 percent increase of hourly wages in West Germany increases the migration rate of youngsters (aged 18 to 24) by 1.4 percent.

Table 9.4 Elasticity of East-West Migration with Respect to East-West Unemployment and Wages

	Age 18–24	Age 25–49	Age 50–64
Destination hourly wage	1.431	1.061	1.889
Source hourly wage	–0.803	–0.750	–0.102
Destination unemployment	–0.131	–0.259	–0.097
Source unemployment	–0.057	0.151	0.279

Hunt also found that the Roy model predicts quite well the skill composition of young migrants but is less successful with older workers and commuters, who have to face a lower up-front cost of migration and hence may react more to absolute, rather than simply to relative, wage differentials. Overall, the stronger effect of wage differentials (over unemployment differentials) on migration seems to be due to the age composition of migration: there are more young than older migrants. Hunt's findings are particularly relevant in assessing the role played by centralized wage agreements in countries, like Italy and Spain, with large regional differences in unemployment and productivity.

Source: Hunt (2006).

9.3.2 Fiscal Effects

Some studies have recently tried to evaluate the net contribution of migrants to fiscal balances in the country of destination. This literature is relevant to evaluate the effects of migration on the labor market that come via labor taxes.

This literature also is largely confined to the United States. Whether immigration increases or decreases fiscal costs for native taxpayers was analyzed by a

report of the National Academy of Sciences for the U.S. Congress. The results of this study are summarized by Smith and Edmonston (1997), who focus on New Jersey and California. Notwithstanding a number of conceptual and data problems, they offer a rather comprehensive overview of the transfers between natives and immigrants. In 1996 the net annual fiscal burden of immigration, considering all transfers at the local and national levels, was estimated to be between \$166 and \$226 per native household. The main reasons for these sizable transfers from natives to immigrants are differences in family structures and income levels: immigrant families have more dependent children who use publicly funded schools, and immigrant households are poor and hence receive more transfers and pay less taxes. Smith and Edmonston (1997) also document large differences in transfers across U.S. states, which can be explained in terms of differences in demographic structures and welfare schemes across states. The age structure and the ethnic composition of immigrants are also very important in the fiscal impact of immigration. An average immigrant household in California receives net benefits of \$4,977 if the members are from Latin America, but contributes \$1,308 if the members are from Europe or Canada.

Less information is available for Europe. Boeri (2008) and Zimmermann (2005) documented that migrants in some countries are overrepresented in the population of recipients of social assistance, housing benefits, and unemployment benefits, while they are less represented than natives among pensioners and recipients of sickness benefits. This overrepresentation can be largely explained by the personal characteristics of migrants that make them more likely to be eligible for social transfers than natives. Indeed, regressions (run over European Community Household Panel [ECHP] data covering the period 1994–2001) rarely detect the presence of “residual dependency from social transfers” by non-EU citizens (Boeri 2008). In other words, the fact of being a migrant is not relevant to determine receipt of benefits once account is taken of the personal characteristics of the individual (e.g., the number of dependent children).

9.4 Policy Issues

The preceding discussion suggests that migration, notably unskilled migration, has the potential to increase wage and income inequality among natives and to be a source of fiscal costs for an immobile native workforce. These externalities can reduce the economic gains of migration and create political opposition to migration. Often governments react to these concerns of public opinion by tightening migration restrictions, as we documented in section 9.1. However, migration restrictions are rarely effective in preventing migration altogether. Often they end up simply increasing illegal migration, which involves mainly unskilled workers, and preventing these workers from having a regular job and hence paying social security contributions. There is a high risk in this context of a vicious circle

being set in motion where unrealistic restrictions to migration induce more illegal and unskilled migration, which strengthens public opinion against migrants and pushes governments to adopt even stricter (and unenforceable) migration restrictions. Two strategies attempt to address this potential vicious circle:

1. Closing the welfare door to migrants
2. Introducing a points system that rewards skilled migration

9.4.1 Closing the Welfare Door?

A number of scholars, including Hans-Werner Sinn (2004a, 2004b), have advocated the temporary closing of the welfare door to migrants. Closing the welfare door would reduce migration flows while mildly increasing the proportion (not the absolute number) of skilled workers in migration inflows. Closing the welfare door, however, would postpone the assimilation of migrants who are already in the country or who would come in any event. Thus it might paradoxically increase negative externalities on natives of immigration to rigid labor markets by pushing many migrants into illegal activities.

Another rather subtle argument often made for closing the welfare door uses political economy considerations. A policy explicitly preventing abuse of welfare by migrants is deemed to buy popular support for more realistic migration policies in individual EU countries. Analysis of opinion surveys (Boeri and Bruecker 2005; Dustmann and Gitz 2005) shows that many citizens are indeed concerned about welfare abuse by migrants. Yet this does not imply that voters would support more liberal policies when migrants are banned from some citizenship rights (e.g., access to social welfare) in the country of destination.

Closing the welfare door may not be a credible policy in countries with large immigration. The U.S. experience is revealing in this respect. In 1996 the welfare system was partly decentralized to the states, and limitations were introduced to access to welfare benefits for legal immigrants. For instance, legal nonasylum immigrants who arrived in the country after August 1996 were barred from receiving food stamps or using Medicaid for five years. The proponents of this reform were hoping that a more decentralized system would make states more cautious in providing expensive welfare benefits to immigrants. The reform failed (Boeri, McCormick, and Hanson 2002). Since 1996 the provision excluding immigrants from some welfare services has been challenged in the courts. In 1997 Congress started repealing the tougher provisions. Finally, the states felt political pressure to maintain benefits at previous levels under the federal system. This pressure is particularly evident in high-migration states like California, where immigrants account for more than 15 percent of the electorate. Thus a decentralized system that strongly discriminates against immigrants can face political resistance, is easily challengeable in the courts, and is likely to ultimately revert to the previous system.

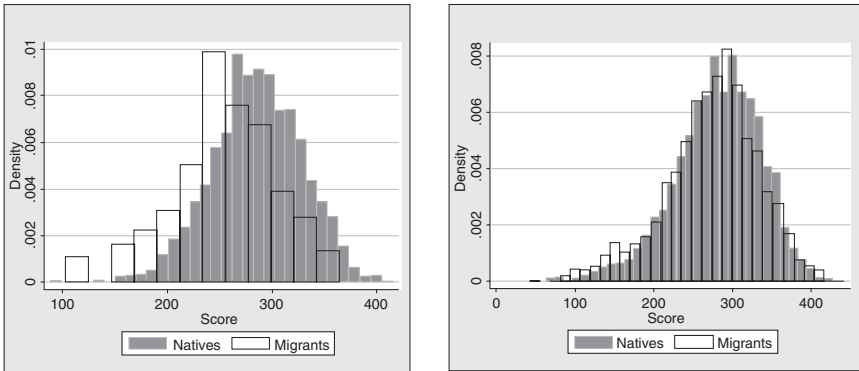


Figure 9.7 Distribution of IALS Test Scores among Natives and Migrants in Germany (left) and Canada (right), 2001

9.4.2 Adopting a Points System?

As discussed in the introduction to this chapter, a points system is a method to rank applications for residence and work permits. It has been adopted by Australia, Canada, New Zealand, and Switzerland, and the United Kingdom and other European countries are in the process of adopting a points system. Each application is allocated a score based on explicit criteria, which typically reward educational attainment, experience, and language abilities. Bonus points can also be given for employment in occupations and regions where there is a shortage of workers. A well-designed points system could encompass and simplify the entire range of migration regulations, for example, removing the need for ad hoc policies for highly skilled migrants and integrating asylum policies into a broader framework. Hatton (2004), for instance, suggested introducing humanitarian points.

Points systems seem to be quite effective in selecting migrants. This can be seen by comparing the distribution of 2004 IALS (International Adult Literacy Survey) scores of migrants versus natives in a country without a points system (Germany) and one with a points system in place, like Canada (figure 9.7). The difference is quite striking both in absolute skill levels (the average IALS score for migrants was roughly 300 in Canada compared with 250 in Germany) and relative to the skills of natives. In Germany the skill distribution of migrants is visibly tilted to the left, while in Canada it almost perfectly overlaps the distribution among natives. Difference-in-differences analysis of changes in the skill composition of migration in Canada after the introduction of the points system and comparison with evolutions in the United States also confirm the effectiveness of the points system. In Canada the fraction of migrants with tertiary education increased by 5 percentage points from 1987 to 2002, while it decreased in the United States. However, the points system may also discriminate against certain countries of origin rather than simply being based on skill levels (Antecol, Cobb-Clark, and Trajo 2003).

As discussed earlier, immigration of skilled workers may reduce income inequalities in the recipient country and help contain fiscal pressures associated with access to the welfare system. Thus it can reconcile efficiency and equity goals of migration policy. Moreover, there is evidence of skill complementarity; hence inflows of skilled migrants may even increase job opportunities of unskilled natives.

There is, however, an important drawback of a points system, its negative effect on growth in the sending country via the *brain-drain* effect. The brain drain, as its name suggests, is the assertion that migration tends to strip the sending nation of all its best workers. Evidence on the brain-drain effect of migration is, however, far from conclusive (Docquier and Rappoport 2004). Selective immigration policies increase individual incentives to invest in human capital in the sending countries, so the impact of migration on human capital formation in the country of origin is ambiguous. Moreover, since migration is often temporary, human capital acquired in the country of destination could subsequently be transferred to the country of origin and promote growth in the sending region. Finally, access to higher rewards for their skills enables migrants to send generous remittances back home, supporting growth there.

9.5 Why Do Migration Policies Exist?

Migration policies are essentially redistributive policies. They typically prevent income inequality from rising at the bottom of the income distribution by sheltering unskilled native workers from competition with foreign workers. Hence these policies are typically supported by workers at the low end of the skill distribution. Immobile taxpayers can also benefit from tight migration policies because they are vulnerable to fiscal spillovers associated with inflows of migrants who draw benefits from social welfare systems. Aging increases the political power of these groups, while an increase in the educational attainments of the domestic workforce may weaken the antimigration constituency. The latter is becoming increasingly important in affecting policy outcomes. As documented in this chapter, the trend seems to be toward a tightening of migration restrictions.

Restricting migration, however, may not be very effective in addressing the concerns of these groups because these restrictions may paradoxically end up selecting just that type of migrants whose assimilation and integration into the fiscal system are more problematic. In any serious policy evaluation there is not a zero-immigration scenario. Migration policies can, at best, induce some gradualism in migration flows that would otherwise occur in large waves, as predicted by the option value theory of migration reviewed in the technical annex. Migration policies, in any event, cannot altogether prevent migration.

Pressures on welfare systems that exert negative fiscal spillovers on the domestic population can be reduced by either restricting access to welfare by migrants or

by adopting a points system. We have weighed pros and cons of these two types of measures.

9.6 Suggestions for Further Reading

A nice history of migration is provided by Tim Hatton and Jeffrey Williamson (1998). The book edited by Boeri, Gordon Hanson, and Barry McCormick (2002) offers surveys of the literature on the effects of migration on the U.S. and European labor markets. George Borjas (2001) provides an explanation of the small effects of immigration on wages and employment found by this literature in terms of self-selection of immigrants, as well as estimates of the so-called greasing-the-wheels effects. Michael Burda and Jennifer Hunt (2001) offer an assessment of the German reunification episode that sheds light on the determinants of migration flows in integrated labor markets. Christian Dustmann and Albrecht Gitz (2005) nicely review evidence from survey data on attitudes of residents about immigration. Tito Boeri and Herbert Bruecker (2005) draw lessons from the eastern enlargement of migration to the European labor market. Finally, David Card (2005) offers an up-to-date assessment of the consequences of migration to the United States for wages of low-skilled natives.

9.7 Review Questions

1. Why do employers generally support migration, while unions do not?
2. What are the effects of migration on income distribution at home?
3. What is the relation between internal and international migration?
4. Why does empirical work often not find the strong effects of migration on native wages that are predicted by economic theory?
5. What does the Roy model predict about the skill composition of migration?
6. How do the elasticities of labor demand and labor supply affect the economic impact of migration?
7. What is the greasing-the-wheels effect?
8. What are the pros and cons of a points system?
9. What are the pros and cons of a policy that closes the welfare door to migrants?

9.8 Technical Annex: Net Gains from Migration and the Option Value of Waiting

The decision to migrate can be modeled just like a human capital investment decision (see chapter 8). Suppose that an individual at the beginning of a career

($t = 0$) has to choose between staying in a given country, say, home (H), or migrating to a foreign country (F). If the individual stays at home, the annual earnings over the remaining labor market years—from $t = 1$ until age T^e —will be $w_H(t)$. If the individual moves from country H to country F , from year 1 onward the individual will earn a stream of income $w_F(t)$. The costs of moving are all incurred at the time of moving and are equal to C_0 . The individual makes a decision on the basis of a comparison of present values of both flows of earnings and costs. If i is the market interest rate, $\delta = \frac{1}{1+i}$ is the monetary discount rate the individual uses. The net present value of staying at home is given by

$$NPV_H = \sum_{t=1}^{T^e} \delta^t w_H(t), \quad (9.2)$$

while the net present value of working in the foreign country since is given by the beginning of the career

$$NPV_F = -C_0 + \sum_{t=1}^{T^e} \delta^t w_F(t). \quad (9.3)$$

The individual will migrate if the present value of working abroad minus the costs of moving is larger than the present value of remaining at home. This condition requires that:

$$\sum_{t=1}^{T^e} \delta^t (w_F(t) - w_H(t)) > C_0. \quad (9.4)$$

From this it follows that the lower the costs of mobility, the younger the individual (the longer the time horizon, T^e , in which the initial investment can provide returns in terms of earning differentials), and the larger the difference in wages between destination and origin countries, the higher the probability that the individual migrates.

Suppose that wages are constant over the lifetime both in the country of origin and in the country of destination. Then for large T^e the net present value of migrating (NPV_M) can be approximated by

$$NPV_M \approx \frac{T^e (w_F - w_H)}{1 - \delta} - C_0,$$

where we have used the properties of geometric series. Dividing both terms by the wage at home, we have that migration occurs if

$$\frac{(w_F - w_H)}{w_H} > \frac{C_0}{w_H} \frac{i}{(1+i)T^e}.$$

This suggests that even if the costs of migration are very large relative to the wage in the home country, say, 20 times as large as a yearly wage, the discount

rate is 5 percent, and the length of the worklife is 40 years, then for migration to be convenient, it is sufficient that wages in the destination country are just 2.5 percent larger in the country of destination than in the country of origin. To put it another way, migration costs must be very large; otherwise we would observe large migration flows even in the presence of relatively modest wage differentials. Another interpretation is that there is uncertainty associated with the stream of incomes in the country of destination, and this uncertainty implies that there is an option value of waiting. The latter explanation is consistent with the observation of large waves of migration flows.

When there is no uncertainty as to future earnings in the country of destination, it is always better to migrate at the beginning of the working life in order to capitalize on the wage differential. Suppose now that there is some uncertainty associated with the future stream of earnings in the country of destination, and that we allow the individual to be able to decide the date, if any, in which to migrate. As explained in technical annex 6.8, there may be now an option value of postponing migration (Burda 1995). Postponing migration implies that the individual will have less time to capitalize on earning differentials. However, the individual may gather more information that allows a better choice of destination country. Suppose that this additional information comes at some date $0 < \tau < T^e$. For instance, it may be that conationals will come back home with information on the labor market in the country of destination.

The option value of waiting to migrate until time τ is given by

$$OV = NPV_M(\tau) - E_0 NPV_M(0).$$

A person will postpone retirement if the option value of waiting to migrate is positive and will migrate as soon as this option value becomes negative. This effect induces large waves in migration flows when the option value becomes negative for a large group of migrants, for example, migrant coming from the same country.

This page intentionally left blank

Employment Protection Legislation

Employment protection legislation (EPL) consists of the set of norms and procedures to be followed in case of dismissals of redundant workers. EPL imposes legal restrictions on dismissals and compensations to workers to be paid by their former employers in case of early termination of a permanent employment contract. A number of procedures are also envisaged under EPL that have to be followed in case of both individual and collective layoffs. The final decision on the legitimacy of a layoff generally depends on a court ruling. From the point of view of economic analysis it is very important to note that the firing decision is not only up to the worker and/or the employer, but also can involve the participation of a court, a third party, which can be requested to assess the legal validity of the layoff.

EPL is a multidimensional institution, although from the standpoint of economic theory it can be reduced to two key components: a tax and a transfer. The *tax component* is a monetary transfer from the employer to the worker, similar in nature to the wage. The *transfer component*, instead, is more like a tax because it corresponds to a payment to a third party, external to the worker-employer relationship. Conceptually, the transfer component of EPL consists of *severance payments* and the mandatory *advance notice period*, while the tax component consists of *trial costs* (the payments for lawyers and the like) and all the other *procedural costs*. *Severance payments* refer to a monetary transfer from the firm to the worker to be paid in case of firm-initiated separation. *Advance notice* refers to a specific period of time to be given to the worker before a firing can actually be implemented. Both the severance payment and advance notice that are part of EPL refer to the *legal minima*, that is, statutory payments and mandatory rules that apply to all employment relationships, regardless of what is established by specific labor contracts. Beyond mandatory payments, collective agreements may specify larger severance payments for firm-initiated separations. Another important dimension of EPL consists of the *administrative procedures* that have to be followed before the layoff can actually take place. In most countries the employer is often required to discuss layoff decisions with workers' representatives. Furthermore, legislative provisions often differ depending on business characteristics such as firm (or plant) size and industry of activity.

In most countries legislation distinguishes between *individual* and *collective* dismissals. Individual dismissals should be further distinguished between *economic* dismissals and *disciplinary* dismissals, with most EPL clauses applying only to the former case. Disciplinary dismissals (i.e., worker's fault dismissals) typically do not involve monetary transfers. The procedure for collective dismissals applies to large-scale firm restructuring and requires the dismissal of at least a specific proportion (often 10 percent or an absolute number, say, 50 workers, for larger firms) of the workforce. When a collective dismissal is authorized by the relevant authority, the firm can then implement large-scale dismissals with lower transfers than would be required by applying the individual layoff provisions to all workers. Yet this procedure requires a much tighter administrative burden and procedural costs, in the form of prolonged consultation with workers' representatives.

EPL bears a close relationship to unemployment benefit systems (UBs, chapter 11) in providing insurance against unemployment risk. EPL protects only those who already have a job, while UBs are typically funded by a payroll tax and protect also the unemployed (at least those with previous work experience). Insofar as EPL involves payments by the employers implementing the layoffs, it internalizes the fiscal externality associated with redundancies (the fact that workers are bound to receive unemployment benefits that are paid by all workers and employers, rather than only those responsible for the redundancies).

10.1 Measures and Cross-Country Comparisons

From a cross-country perspective, it would be better to measure EPL by evaluating the average cost of a layoff under different employment protection regimes. Unfortunately, homogeneous measures of these costs relative to the average wage do not exist for all countries. In order to carry out international comparisons of employment protection regimes, economists use the so-called method of the hierarchies of the hierarchies. This method amounts to assigning a number (say, from 1 to 6) to every country for any single feature of the protection regimes. Higher numbers denote more rigid regimes. By taking the average of the several components, a single synthetic measure of the strictness of EPL is then obtained. The synthetic indicators, originally compiled by the OECD, are now available for a relatively large set of countries. In order to obtain the overall indicator of the rigidity of a country, it is necessary to consider simultaneously (1) the rigidity of the firing regulations for individual workers under *permanent contract*, (2) this rigidity for workers under *temporary contracts*, and (3) the rigidity of *collective dismissals*. The average of these three measures gives the overall indicator.

Obviously, each of the three measures is obtained, in turn, through an average of some other submeasures. As an example, consider the indicator for individual firings of workers under permanent contracts. This number is obtained as an average of 4 subindicators: (1) the administrative procedures, (2) the length of the advance

notice period, (3) the amount of the severance payment, and (4) the severity of enforcement (the more or less important role of judges in firing disputes).

A number of caveats apply to the interpretation of the OECD EPL index. First, the OECD overall index averages a number of subindicators covering regular employment contracts, temporary contracts, and collective dismissals. The overall index can therefore change because of variations in one or more subindicators. For this reason it is always advisable to look at the subindicators and use them in the econometric analysis rather than concentrating on an aggregate index.

Second, changes in the EPL subindicators are not independent of one another. For instance, the increasing shares of employment under fixed-term contracts may also be a consequence of strict employment protection for regular workers because more investment may be directed to activities exempted from these provisions, rather than an indication of greater labor market flexibility per se (Bertola et al. 2000). In other words, the coverage of EPL for regular workers is endogenous. As documented later, the share of employment exempted from these regulations (e.g., the share of employment with temporary contracts, the size of the informal sector, the fraction of workers in small units) is larger in countries with particularly strict EPL regimes for permanent contracts. This suggests that it is always better to weight the OECD index for regular workers by the share of workers subject to these regulations. This is done later in table 10.2.

Third, ideally one should obtain indicators of the two crucial components of EPL from the standpoint of economic analysis: the *transfer* and the *tax* components. However, disentangling the tax from the transfer component is not an easy task because it is particularly difficult to measure the legal costs of dismissals, whose amount depends on the probability that the worker files the case in a court, and on the probability that the court invalidates the firm's firing. To give an example, in the Italian case, when the firing decision is overruled by a judge, the firm can be forced to reinstate the employee on the payroll. Garibaldi and Violante (2005) estimated that an Italian employer with more than 15 employees who fires a worker and whose decision is overruled by a court a year after the layoff with an 80 percent probability will have to bear a cost of 15 monthly wages, that is, a year and three months of wages. This amounts to roughly 20 percent of the total costs of the layoff. The distinction between the tax and transfer components of EPL can, however, be approximated by attributing to the tax component the administrative procedures and severity of enforcement measures and to the transfer component the subindicators referred to severance payments and advance notice periods.

With these caveats in mind, table 10.1 displays the overall EPL index and its three subcomponents at the end of the 1980s and in 2003, the last year for which data are available. Two facts are noteworthy. First, there are marked differences across countries in the strictness of EPL. The United States turns out to be the most flexible country, while EPL is much tighter in countries such as Portugal and the new EU member states. Second, EPL for regular workers (workers with permanent

Table 10.1 Strictness of Employment Protection: OECD Index

	Overall		Regular		Temporary		Collective	
	Late 1980s	2003	Late 1980s	2003	Late 1980s	2003	Late 1980s	2003
Australia	0.9	1.2	1.0	1.5	0.9	0.9	2.9	2.9
Austria	2.2	1.9	2.9	2.4	1.5	1.5	3.3	3.3
Belgium	3.2	2.2	1.7	1.7	4.6	2.6	4.1	4.1
Canada	0.8	0.8	1.3	1.3	0.3	0.3	2.9	2.9
Czech Republic	–	1.9	–	3.3	–	0.5	2.1	2.1
Denmark	2.3	1.4	1.5	1.5	3.1	1.4	3.9	3.9
Finland	2.3	2.0	2.8	2.2	1.9	1.9	2.6	2.6
France	2.7	3.0	2.3	2.5	3.1	3.6	2.1	2.1
Germany	3.2	2.2	2.6	2.7	3.8	1.8	3.5	3.8
Greece	3.6	2.8	2.5	2.4	4.8	3.3	3.3	3.3
Hungary	–	1.5	–	1.9	–	1.1	2.9	2.9
Ireland	0.9	1.1	1.6	1.6	0.3	0.6	2.4	2.4
Italy	3.6	1.9	1.8	1.8	5.4	2.1	4.9	4.9
Japan	2.1	1.8	2.4	2.4	1.8	1.3	1.5	1.5
Korea	–	2.0	–	2.4	–	1.7	1.9	1.9
Mexico	–	3.1	–	2.3	–	4.0	3.8	3.8
Netherlands	2.7	2.1	3.1	3.1	2.4	1.2	3.0	3.0
New Zealand	–	1.5	–	1.7	–	1.3	0.4	0.4
Norway	2.9	2.6	2.3	2.3	3.5	2.9	2.9	2.9
Poland	–	1.7	–	2.2	–	1.3	4.1	4.1
Portugal	4.1	3.5	4.8	4.3	3.4	2.8	3.6	3.6
Slovak Republic	–	1.9	–	3.5	–	0.4	3.3	2.5
Spain	3.8	3.1	3.9	2.6	3.8	3.5	3.1	3.1
Sweden	3.5	2.2	2.9	2.9	4.1	1.6	4.5	4.5
Switzerland	1.1	1.1	1.2	1.2	1.1	1.1	3.9	3.9
Turkey	–	3.7	–	2.6	–	4.9	1.6	2.4
United Kingdom	0.6	0.7	0.9	1.1	0.3	0.4	2.9	2.9
United States	0.2	0.2	0.2	0.2	0.3	0.3	2.9	2.9

Source: OECD (2004b).

Notes: Higher numbers denote stricter EPL regimes. All subindexes are normalized in the 0–6 range.

contracts) hardly changed at all in OECD countries over the period covered by the data. Conversely, the regulation of temporary contracts was eased in most European countries. In particular, the scope of fixed-term contracts was significantly expanded, and temporary work agency (TWA) was introduced, allowing firms to face temporary peaks in demand without having to hire new workers permanently. This corresponds to a *dual-track* reform strategy, involving reforms only at the margin, on a flow basis, for new hires, while the employment security entitlements of the incumbent workers remain unchanged.

In order to gauge the importance of these dual-track reform strategies in reducing the scope of the most restrictive EPL provisions, table 10.2 presents an adjusted

Table 10.2 Adjusting EPL by Coverage

Country	EPL	Coverage %	Coverage	(a)*(b)	(a)*(c)
	Regular worker (a)		(net of shadow employment) (c)		
Australia	1.5	75.6	70.7	1.13	1.06
Austria	2.9	79.6	66.9	2.31	1.94
Belgium	1.7	77.7	—	1.32	—
Canada	1.3	74.8	71.5	0.97	0.93
Czech Republic	3.3	77.5	—	2.56	—
Denmark	1.5	83.3	70.5	1.25	1.06
Finland	2.2	73.3	—	1.61	—
France	2.5	76.7	69.8	1.92	1.74
Germany	2.7	76.0	61.5	2.05	1.66
Greece	2.3	56.9	—	1.31	—
Hungary	1.9	80.7	—	1.53	—
Ireland	1.6	82.1	—	1.31	—
Italy	1.8	65.1	45.6	1.17	0.82
Japan	2.4	78.6	—	1.89	—
Netherlands	3.1	76.8	—	2.38	—
Norway	2.3	85.4	—	1.96	—
Poland	2.2	56.4	—	1.24	—
Portugal	4.3	64.0	—	2.75	—
Slovak Republic	3.6	83.0	—	2.99	—
Spain	2.6	54.7	42.7	1.42	1.11
Sweden	2.9	76.0	60.9	2.20	1.77
Switzerland	1.2	74.7	—	0.90	—
United Kingdom	1.1	83.2	—	0.92	—
United States	0.2	86.7	—	0.17	—

Sources: Estimates based on OECD LFS Database (2006) and Schneider and Enste (2000).

Notes: (a) OECD EPL index for regular workers (see table 10.1).

(b) = 100 – (share of fixed-term contract in total employment) – (share of self-employment in total employment).

(c) = (b) × [1 – (shadow employment rate)].

index for permanent contracts that takes into account also the coverage of these provisions. In particular, the coverage is calculated by estimating the share of permanent workers in the economy. The latter is obtained by excluding from total employment workers employed with a temporary contract (either part-time or full-time), as well as workers who are self-employed, and by using, whenever available, estimates of shadow employment (Schneider and Enste 2000), that is, workers employed in the informal sector.

As shown by the table, once the actual coverage of these regulations is taken into account, the OECD ranking of countries by strictness of EPL is quite substantially altered. In particular, southern European countries appear to be much less strict than in the OECD index because they have a large share of employment involved in temporary contracts and a sizable informal sector.

In addition to temporary contracts, there are many other exemptions from the application of EPL rules on permanent contracts. For instance, small units are usually exempt from the reinstatement obligations and other procedural requirements in the light of the fact that these obligations are a source of fixed costs and hence particularly affect small units. The relatively small size of plants and the large informal sector in southern Europe can also be a by-product of the presence in these countries of very rigid regimes for regular contracts.

10.2 Theory

EPL imposes costs on firms for adjustments of the levels of the workforce. To put it another way, employers can avoid paying severance payments and procedural costs by deciding not to change employment levels in response to shocks. Lack of response to shocks could reduce profits of firms, but this loss could possibly be compensated by a decline in wages if employers succeed in shifting to workers the EPL tax as a sort of voluntary insurance scheme. This suggests that in order to fully characterize the effects of EPL, we need a framework with (1) labor market flows, notably job creation and job destruction, and (2) endogenous wage determination. This framework is provided in technical annex 10.8. Because the model is fairly complicated, we will go as far as we can by using a simple static framework and, in any event, confine ourselves here to presenting some numerical simulations from the dynamic model, describing the adjustment mechanism involved under various assumptions about wage determination and the nature of EPL.

10.2.1 A Neutrality Result

Under some rather extreme circumstances EPL has no effects on employment, welfare of workers, and profits. This neutrality case was first highlighted by Edward Lazear (1990) and occurs when three conditions are met:

1. Workers are risk neutral.
2. Wages are flexible; that is, there is no wage floor (minimum wage or collective agreement) preventing downward wage adjustment.
3. EPL consists only of the transfer component, a severance payment to be paid by the employer to the worker involved in the redundancy; the tax component is zero.

Neutrality occurs under these circumstances because EPL affects only the intertemporal structure of wages, leaving the net discounted value of a job for a worker and the worker's employer unchanged. Actually, the employer may use the severance payment as a device to induce the worker to remain with the firm until the end of the contract.

The essence of this result is that the presence of a severance payment is taken into account in the wage contract, which internalizes the future redundancy payment. Employers, in other words, initially pay a lower wage, forcing their employees to buy from them a sort of *bond* or insurance that will give them the right to receive a deferred compensation, the severance payment, at the time of separation, as in the example discussed in box 10.1.

Box 10.1 Example of a Contract Undoing EPL

Suppose we have initially a two-period contract offering wages w in both periods. Under the assumptions of the Lazear model, workers are risk neutral, so their (indirect) utility function reads $u(w) = w$, and there are no wage floors. Now introduce employment protection as a severance scheme, paying TR to the worker at the termination of the two-period contract. In order to keep labor costs unchanged with respect to the situation without EPL, the employer proposes to the worker a contract offering a lower wage in the first period, $(w - B)$, where B is the *bond* entitling the worker to TR in the second period. For the worker, there is no loss in welfare as long as

$$w - B + \frac{w + TR}{1 + i} = w + \frac{w}{1 + i},$$

where i is the interest rate. This condition implies that $B = \frac{TR}{1+i}$, that is, the worker is lending to the firm the transfer that will be delivered at the end of the contract, as depicted in figure 10.1.

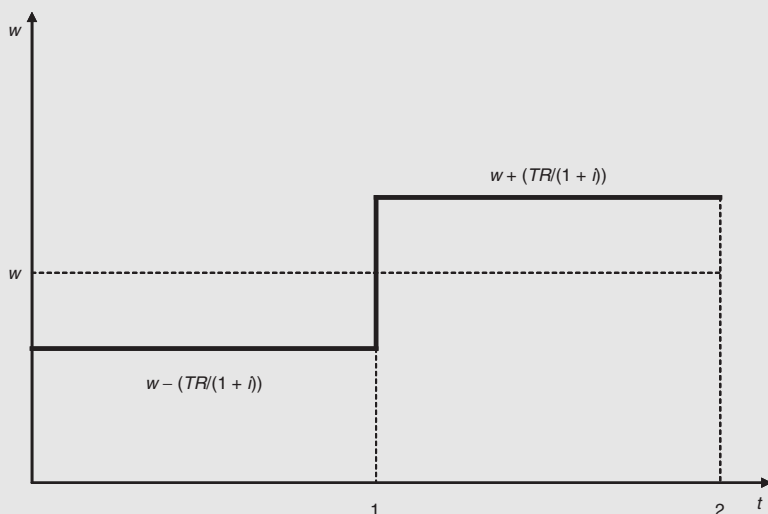


Figure 10.1 A Bonding Scheme Undoing the Severance Payment

(continued)

Box 10.1 (*continued*)

The risk-neutrality assumption is crucial in this context. Suppose in fact that the utility function $u(w)$ is concave, denoting a risk-averse worker. It follows that

$$u(w) > \frac{1}{2}[u(w - B) + u(w + B)].$$

Hence

$$u(w) + \frac{u(w)}{1+i} > u\left(w - \frac{TR}{1+i}\right) + \frac{u(w + TR)}{1+i}.$$

Thus a bonding arrangements will cause a welfare loss to the employees under these circumstances.

Source: Lazear (1990).

The key to the neutrality result is that the worker receives the same payment in discounted value terms from offering labor with or without EPL. A risk-neutral worker is interested only in the discounted value of the job, not in the time profile of the wage, and the employer, when deciding whether or not to create a job and open a vacancy, will be indifferent to the presence or not of EPL because the net discounted cost of the job is unchanged with respect to the situation without EPL. The contract has succeeded in undoing EPL, which does not affect either labor demand on labor supply.

10.2.2 Removing Risk Neutrality

It is sufficient to relax any of these three assumptions to have some effects on EPL on labor allocation. Take first the case in which workers are risk averse (their indirect utility function is concave rather than linear in wages). This means that they will suffer a welfare loss from experiencing fluctuations in their earnings even if the net discounted value of the job is unchanged with respect to the situation without EPL (see box 10.1).

Workers may then ask for higher wages in order to be compensated for the fluctuations in earnings associated with the bonding arrangement. As labor supply shifts upward the bonding scheme cannot be neutral for the employers as well: the introduction of EPL is bound to affect the labor market equilibrium.

10.2.3 EPL with Rigid Wages

Suppose now that assumption 2 is relaxed, that is, that wages are rigid and therefore do not adjust after the introduction of EPL. Because the introduction

of the severance payment cannot be accompanied by a reduction of the wage, labor supply will be unaffected, and we can concentrate on labor demand, as in box 10.2.

Box 10.2 Flexiland and Rigidland

In order to highlight the effects of EPL under rigid wages, it is instructive to consider an economy where only two states of the world are possible—a good and a bad state, for example, a boom and a recession—affecting the value of the marginal product of the firm. In particular, let us take a simple logarithmic production function (labor is the only factor of production) and hence write firm's profits in the *flexible regime* (denoted with superscript F) as follows:

$$\pi^F = A^i \log L - wL,$$

where L is labor and A^i is the price of the good being sold by the firm, which varies depending on the state of the world: A^i assumes value A^h under the good state and A^L under the bad state, where $A^h > A^L$. In every period there is a probability p that the price is equal to A^h and a probability $1 - p$ that the price is equal to A^L . The wage is fixed and equal to w independently of the cyclical conditions.

The employer has to decide how many workers to hire. Without EPL, hiring and firing can take place at no cost, and the firm can freely choose the profit-maximizing employment level (equating the value of the marginal product of labor to its marginal costs or wage rate) under any state of the world. With EPL, the firm is instead forced to keep always the same employment level regardless of cyclical conditions; we are here assuming that EPL is unboundedly expensive, so that an employer would never lay off a worker.

Without EPL, the firm chooses employment after observing the productivity level. Optimal employment in this case equates the value of the marginal product, $\frac{A^i}{L}$, to the wage, so that, solving for employment,

$$\begin{aligned} L^F &= \frac{A^L}{w} & \text{if } A^i = A^L, \\ L^F &= \frac{A^h}{w} & \text{if } A^i = A^h. \end{aligned}$$

Thus the firm will fire (hire) $\Delta L^F = \frac{A^h - A^L}{w}$ workers when the economy moves from booms (recessions) to recessions (booms), as depicted in figure 10.2. Because the economy experiences, on average, a fraction p of booms and a fraction $(1 - p)$ of recessions, the average employment in the long run will be

$$\bar{L}^F = \frac{(1 - p)A^L + pA^h}{w}.$$

Consider now the behavior of the firm when EPL is present. Because laying off workers during recessions is too expensive, the firm can only choose the

(continued)

Box 10.2 (continued)

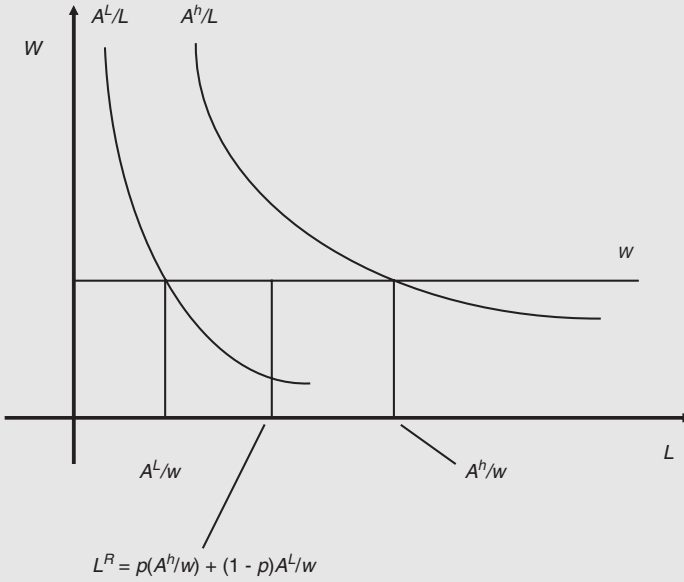


Figure 10.2 Employment in the Flexible and Rigid Regimes

employment level that maximizes the *expected* value of the profits, holding this level constant over time, independently of cyclical conditions. This implies that the employer in the *rigid regime*, R , will solve the expected profit-maximization problem

$$\pi^R = \max_L \{[(1-p)A^L + pA^h] \log L - wL\}.$$

The first-order conditions of this problem provide the value of employment in the rigid regime as

$$L^R = \frac{(1-p)A^L + pA^h}{w}. \quad (10.1)$$

This optimal employment level in the rigid regime, L^R , is therefore a weighted average of the levels of employment that, without EPL, would prevail during the expansions and recessions, respectively, where weights are given by the probabilities of the two events. Notice that L^R coincides with $\bar{L}^F$, the long-run employment level in flexiland. Thus employment in the long run is the same in the two regimes. However, during the cycle, employment in the rigid regime will never be as in flexiland: it will be lower under upturns and higher during downturns. The only case when L^R equals L^F (employment in flexiland in the short run) is when the economy is always in one of the two states of the world (if $p = 0$ or $p = 1$). Finally, notice that $\pi^R < \pi^F$ even in the long run because the firm realizes lower profits (using suboptimal employment levels) in both booms and recessions.

Generalizing from the results in box 10.2, we conclude that EPL with fixed wages

1. has no effects on average employment or unemployment,
2. lowers the volatility of employment over the business cycle, and
3. reduces profits.

The first two implications are intuitive in the light of the preceding discussion. The third draws from the fact that the employment level chosen without EPL is the only level that maximizes profits in each period. Consequently, in each period profits are higher without EPL than with EPL. With the same level of employment on average, the firm without EPL is able to make a higher level of profits. In other words, the firm operating without EPL is more efficient.

10.2.4 EPL as a Tax

Finally, relax assumption 3 of the Lazear model and let EPL no longer be a transfer, but a tax, notably a payment to a third party, say, a lawyer. Now, even when wages are flexible and workers are risk neutral, EPL can no longer be undone by a new labor contract compensating the employer for the tax.

The effects of employment protection on employment and wages when EPL is a tax can be characterized only by using a dynamic framework like the model developed in technical annex 10.8. Dynamics are needed because the EPL tax is a particular type of tax on the firm, namely, a tax that the employer can avoid paying if it does not reduce employment levels over time. This means that the presence of the EPL tax is taken into account by the employer also when issuing a vacancy and hiring an employee because this hiring carries with it a higher risk of having to pay the EPL tax in the future. To put it another way, the EPL tax displays its effects by acting mainly on labor market flows, on hiring and separation rates.

As shown in the technical annex, an EPL tax reduces job creation because employers are more reluctant to open a vacancy: the net discounted value of a job is lower with EPL than without EPL. At the same time, however, job destruction is also lower in the presence of EPL because it is more costly for firms to lay off workers. The effects of EPL on employment and unemployment are therefore ambiguous; they can increase or decrease depending on the relative strength of the effects on job creation and job destruction margins. An unambiguous theoretical prediction, however, is that labor market flows decline with employment protection: a world with an EPL tax is one with less job creation, job destruction, and unemployment inflows, as well as unemployment outflows, and hence longer unemployment duration (Bentolila and Bertola, 1990).

As long as wages are flexible, they will also react to the introduction of EPL. As argued earlier, an EPL tax cannot be undone by contractual arrangements between

the employer and the worker. Yet wages are nevertheless bound to be affected by EPL. Suppose, for instance, that wages are set as a result of a bargaining process between the employer and the worker, as in the models discussed in chapter 3. These models predict that wages are increasing with the bargaining power of workers and decreasing with their fallback option, that is, the net discounted value of being unemployed. EPL has two effects on wage setting in this context. On the one hand, it increases the bargaining power of those who have a job, the *insiders*, who now are more protected from wage underbidding by unemployed jobseekers, or *outsiders*. This increases the equilibrium wage in comparison with the situation without EPL. On the other hand, EPL reduces the fallback option of workers because unemployed jobseekers face lower reemployment probabilities under strict EPL. This second effect reduces wages under the EPL tax regime. Because of these two offsetting effects, the relationship between EPL and wages is likely to be hump shaped: for low values of EPL, wages are bound to increase as a result of increased bargaining power of insiders, but for high values of EPL, wages may well decline because the welfare loss associated with unemployment becomes very large. Figure 10.3 displays numerical simulations of the model in technical annex 10.8 that correspond precisely to this nonmonotonic profile of the relationship between wages and EPL: wage levels are expressed as fractions of monthly value added.

Even when EPL does not increase wages, insiders are better off with the EPL tax because the risk of job loss is lower. With regard to income distribution, EPL involves lower profits for firms and lower welfare for unemployed individuals, who now experience longer unemployment duration, but higher welfare for those who have a job, notably when individuals care much about their future (have a high discount factor).

10.2.5 Two-Tier Regimes

The preceding theoretical results on the effects of EPL are relevant to predict labor market adjustment in environments with different degrees of strictness of EPL involving all workers. However, as documented earlier, many EPL reforms are asymmetric in that they change regulations only for a subset of the eligible population. This unbundling of reforms, reducing EPL only at the margin for new hirings, seems to be a viable political economy strategy when there are strong political obstacles to reforms (Saint-Paul 1997).

As suggested by Boeri and Garibaldi (2007), these two-tier reforms involve important transitory job creation (*honeymoon*) effects and a decline in productivity. The intuition runs as follows. Starting from a rigid environment where employers do not adjust employment to cyclical conditions (like rigidland in box 10.2), the firm is allowed to enjoy *flexibility at the margin*; that is, it can hire and fire workers on a temporary basis, but at the same time it cannot reduce at will the existing stock

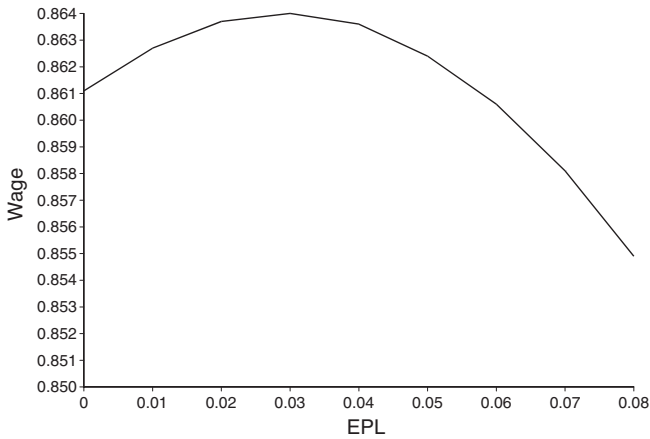


Figure 10.3 Numerical Simulations of the Model in Technical Annex 10.8

of permanent contracts. Thus in good times the firm hires temporary workers up to the optimal employment level in the frictionless environment and dismisses these workers in bad times. This means that during upturns there will be more employees than before the two-tier reforms, while during downturns employment will be as in the fully rigid environment, that is, higher than in the flexible environment. It follows that these dual-track reforms temporarily increase average employment.

These transient honeymoon effects fade away as the stock of permanent workers is gradually replaced with flexible contracts. Natural attrition (retirement) reduces gradually the stock of permanent workers, and hence the firms can build up a buffer stock of workers with flexible contracts (Bentolila and Dolado 1994), allowing them to adjust employment optimally to business conditions. When all workers are replaced, the economy is under a fully flexible regime where the previous theoretical results apply.

Throughout the honeymoon firms enjoy higher profits and lower labor productivity because total output is the same as in the rigid environment, but employment is higher. Profits increase throughout the transition until they reach the same levels as in the flexible regime. The positive effects on employment and the negative effects on labor productivity of two-tier regimes are transitory: they fade away when the honeymoon is over.

10.3 Empirical Evidence

10.3.1 Cross-Country Analyses

Most studies of the effects of EPL take a cross-country approach; that is, they compare employment dynamics in countries that, according to some measure

Table 10.3 The Effects of Employment Protection
on the Labor Market: Empirical Results

Author(s)	Stocks		Flows	
	Employment	Unemployment	Employment	Unemployment
Emerson (1988)	?	?	—	—
Bertola (1990)	?	?	?	—
Lazear (1990)	—	+		
Grubb and Wells (1993)	—			
Garibaldi, Koenig, and Pissarides (1994)	?	?	?	—
Addison and Grosso (1996)	?	?		
Jackman, Layard, and Nickell (1996)	?	?	—	—
Gregg and Manning (1997)	?	?		—
Boeri (1999)	?	?	+	—
Di Tella and McCulloch (1998)	—	+		
OECD (1998)	?	?	?	—
Kugler and StPaul (2000)			+	—
Belot and van Ours (2001)		—		
Nickell, Nunziata, and Ochel (2005)	?	?		

of the strictness of EPL, display different degrees of employment protection. Table 10.3 reviews this literature. A few studies found significant effects of employment protection (generally measured using the OECD cross-country ranking) on employment and unemployment *stocks*. This finding is consistent with economic theory: as a tax on labor adjustment (or on the capitalized value of the firm), EPL should affect hirings and separations, and hence unemployment inflows and outflows, but it should not be correlated with employment and unemployment stocks. Consistent with economic theory, a robust finding of this literature is that EPL negatively affects unemployment inflows and outflows: countries with the strictest EPL have more stagnant unemployment pools. EPL seems also to affect the composition of employment and unemployment (OECD 2004b): countries with more EPL display higher youth unemployment rates and lower unemployment among central age groups. This finding is also consistent with economic theory because central age groups are typically insiders (they already have a job and are sheltered by EPL from labor market adjustment), while youngsters are typically outsiders (they are first-time jobseekers and hence are hit by the longer duration of unemployment).

No unambiguous result is instead obtained concerning the impact of EPL on labor turnover (the sum of hirings and separations) and job turnover (the sum of

job creation and destruction), although economic theory unambiguously predicts a negative effect of the strictness of employment protection on labor market flows. Explanations of this discrepancy between theory and facts—for example, Bertola and Rogerson (1997) and Boeri (1996)—typically call into play the interaction of EPL with other institutional features (chapter 13), as well as measurement errors. For instance, it is argued that institutions that compress wage structures tend to counteract the negative effects of EPL on labor market flows because they reduce the scope of price-driven adjustment mechanisms. If employers cannot adjust wages when they face changes in the demand for their product, they are forced to adjust employment.

These potential interactions with other institutional features question the relevance of a cross-country approach relying mostly on pairwise correlations of country scores in the EPL index and measures of aggregate labor market outcomes. As discussed earlier, there is also quite substantial within-country variation in the actual enforcement of regulations, which is not captured by cross-country analyses.

10.3.2 Within-Country Studies

More and more empirical work is using data referred to the same country and is exploiting any time series available for regulations. Work carried out in the United States has been drawing on cross-state differences in the adoption of wrongful-discharge protections over and above the doctrine of employment at will. Work carried out in Europe has been exploiting another dimension of within-country variation, the exemption of small units from the strictest EPL provisions. By combining these exemptions with dual-track reform strategies, it is possible to carry out difference-in-differences policy evaluation studies. For instance, Boeri and Jimeno (2005) used the difference between permanent fewer than and workers and temporary contracts introduced by dual-track regimes in combination with the difference involved by exemption of small firms (see box 10.3).

Box 10.3 Learning from Within-Country Variation

Boeri and Jimeno (2005) analyzed dismissal rates for temporary and permanent employees in establishments of different sizes by drawing on data from the Italian Labor Force Survey in the 1993–95 period. In Italy the most restrictive EPL provisions concern permanent workers in firms with more than 15 employees. The table in this box obtains an indication of the effects of EPL by drawing on two differences: the difference between workers in firms of fewer than and more than 15 employees, and the difference between permanent workers and temporary workers.

(continued)

Box 10.3 (continued)

	Probability of being dismissed (%)		
	Permanent workers	Temporary workers	Difference
Firms with fewer than 15 employees	1.7	0.8	0.9
Firms with more than 15 employees	0.9	2.2	−1.3
Difference	0.8	−1.4	2.2

Because temporary workers are not covered by EPL provisions upon termination of their contract, their dismissal rates are expected to be higher than those of permanent employees when the strictest EPL provisions for permanent workers hold. This is confirmed by the data in the table: while below the 15-employees threshold the dismissal rate of permanent workers is higher than that of temporary workers, above the 15-employees threshold it is the other way around. These difference-in-differences effects of EPL hold also in different partitions of the sample, for example, by industry, region, gender, age, and skills, and when all these personal characteristics are simultaneously taken into account. Boeri and Jimeno also carried out another difference-in-differences estimation by comparing growth rates of establishments below and above the 15-employees threshold, before and after a reform that in 1990 increased the costs of dismissals for units with fewer than 15 employees. They found that firm size became more persistent after the reform only in units with fewer than 15 employees, an indication of the role played by EPL in reducing adjustment of employment levels in firms. A Stata data file with the Boeri and Jimeno dataset, a do file, and a log file are available at the website <http://press.princeton.edu/titles/8771.html>.

Source: Boeri and Jimeno (2005).

These within-country studies generally found some negative effects of EPL on dismissal rates (Miles 2000; Autor, Donohue, and Schwed 2006; Boeri and Jimeno 2005). Garibaldi, Bargarello, and Pacelli (2003) also found that the presence of firm-size thresholds (e.g. 15 employees, as in Italy) below which EPL does not hold increases firms’ persistence, that is, the probability that a firm does not change the number of employees from one year to the next, just below the threshold. These effects are generally small but qualitatively consistent with the predictions of economic theory.

10.3.3 Endogeneity of EPL

An important issue in the empirical assessment of the effects of EPL is whether the role played by judges in the enforcement of EPL makes these regulations

dependent on underlying labor market conditions. Judges may feel that under severe labor market slack in a region or during a cyclical downturn, workers should be more heavily protected against dismissals than in a buoyant labor market.

Available information on the enforcement of EPL, discussed in Bertola et al. (1999), suggests that the nature and stringency of EPL enforcement do indeed vary across countries and over time and are quite strongly influenced by underlying labor market conditions. In particular, long-term unemployment is positively correlated with the number of cases brought before courts (as a fraction of the population of working age) across the OECD countries for which data on EPL enforcement were available. Another dimension of jurisprudence that seems to be linked to labor market conditions, both across countries and over time, is the percentage of cases favorable to workers: the countries where tribunals are most frequently involved in labor disputes on the termination of a contract are those with the highest percentage of cases favorable to workers. Spain is a case in point. Here almost 72 percent of cases in 1995 were won by workers, compared with less than 50 percent in North American countries and a low of 16 percent in Ireland, all countries where tribunals seem to intervene rather infrequently in labor disputes concerning contract termination. The high incidence of judicial procedures in France may also be partly explained by a large share (74 percent) of cases favorable to workers. The likelihood that court rulings are favorable to employees tends to play an important role in inducing workers to bring their cases to the courts, although it may, on the other hand, also encourage employers to reach extrajudicial agreements.

With regard to time-series variation in EPL enforcement, Bertola et al. (1999) found a marked covariation of the incidence of jurisprudence (cases brought to court as a percentage of the labor force) and unemployment in Germany. Comovements of indicators of jurisprudence and unemployment are also observed in Spain, especially when the focus is on cases ended with sentences favorable to employees, and in Italy, where evidence on case law points to a strong link between law enforcement and regional labor market conditions.

10.4 Policy Issues

10.4.1 How Much Protection Should EPL Provide?

Workers are generally risk averse and have limited, if any, access to capital markets. Their (indirect) utility function is concave in wages. This means that, when comparing two jobs that offer, on average, the same wage, they will always opt for the one that has less variability in earnings. Employers are instead generally risk neutral: firms have better access than workers to capital markets. This means that

they can insure against negative shocks to their business by investing in capital markets and diversifying risk.

Under these conditions it is always optimal to have employers providing some insurance to their workers, allowing them to smooth out income fluctuations. As argued in this chapter, EPL is the most common way for employers to provide such insurance. Unemployment benefits, discussed in the next chapter, offer an alternative. But if we suppose that there is no unemployment benefit system, how much insurance should be provided? Can it be made less distortional?

Blanchard and Tirole (2004) consider a simple case where wages are unaffected by employment protection, there is no asymmetric information (hence workers' effort can be fully monitored by the employer), and EPL consists only of transfers: it is a pure severance scheme. They show that under these conditions the optimal insurance offered by EPL may be a *complete insurance*, preventing any fluctuation in the income of the workers, provided that productivity exceeds the reservation wage of individuals.

However, wages do react to the provision of this insurance. Under rather general circumstances (e.g., a bilateral monopoly in wage bargaining), the theories previously reviewed and the model in the technical annex suggest that EPL has two offsetting effects on wages. On the one hand, EPL strengthens the bargaining power of insiders, exerting an upward pressure on wages. On the other hand, EPL reduces the fallback option of workers, exerting a moderating effect on wages. When workers are fully insured against the risk of job loss (e.g., the severance payment offered by employers compensates them throughout their entire unemployment spell), there is no longer this second, moderating effect on wages related to the outside option of workers. This means that wages are bound to increase unambiguously as a result of EPL, and hence firms providing full insurance to their workers will not be able to partly recover the costs of this insurance through lower wages paid to their workers. If employers are forced by regulations to provide this full insurance, employment will be lower under full insurance.

The presence of moral hazard also prevents the provision of full insurance. Workers require incentives, in terms of a penalty associated with the fact of being fired because of misconduct, in order to avoid shirking, as highlighted by the efficiency wage models summarized in box 10.4. If EPL makes disciplinary layoffs more difficult, then opportunistic behavior is encouraged.

Thus under general circumstances there will be some optimal level of employment protection providing less than full insurance. How much insurance should be provided depends on a number of factors, including the degree of risk aversion of employees, the depth of capital markets, and the underlying wage-setting mechanism. Interactions with other institutions, notably unemployment benefits (chapters 11 and 13), are also very important in this context.

Box 10.4 Efficiency Wages and Employment Protection

The efficiency wage model (Shapiro and Stiglitz 1984) can generate involuntary unemployment as an equilibrium phenomenon. It is also particularly useful in characterizing the effects of employment protection legislation on the productivity of insiders. Unemployment is generated in these models because employers find it optimal to pay a wage above the market-clearing level and use involuntary unemployment as a disciplining device to induce their workers to put more effort into the job. This happens because employers have only imperfect information about the effort of their employees. Hence they pay *efficiency wages* above the market-clearing level in order to discourage their workers from shirking.

The workers face a trade-off. Either they put in effort, which provides a disutility to them, or they shirk and face some positive probability of being detected and being fired. We can use the dynamic framework presented in technical annex 10.8 to illustrate the efficiency wage model in more detail. The value of a job to a person who does not shirk is

$$\rho V_e^N = w - e + \delta(V_u - V_e^N), \quad (10.2)$$

where w is the wage, e is the effort, δ is the exogenous job separation rate (independent of the behavior of the employee), and V_u is the value of being unemployed. For shirkers, the flow value of a job is

$$\rho V_e^S = w + (\delta + q)(V_u - V_e^S), \quad (10.3)$$

where q is the rate at which shirkers are detected and fired, that is, the probability of a disciplinary layoff. The no-shirking condition implies that the value of shirking is lower than or equal to the value of not shirking, that is, $V_e^S \leq V_e^N$. Using this condition together with (10.2) and (10.3), we can then derive the efficiency wage:

$$w^* \geq \rho V_u + e \left(1 + \frac{\rho + \delta}{q} \right). \quad (10.4)$$

From this it appears that the efficiency wage is increasing with the effort required in the job, with the discount rate, and with the exogenous job separation rate, while it is decreasing with the probability of a disciplinary layoff. These effects are intuitively clear. On the one hand, a higher effort, a higher discount rate, and a higher exogenous job separation rate all reduce the value of the job, which means that the penalty associated with being caught shirking, the fact of being laid off by the employer, is lower. Hence there is a need to pay higher wages in order to prevent shirking. On the other hand, a higher probability of disciplinary layoffs reduces the value of the shirking option compared with the value of non shirking, making it less necessary for employers to pay higher wages in order to deter opportunistic behavior.

Employment protection negatively affects both the economic (exogenous) layoff and the disciplinary layoff probabilities. This happens via the costs of

(continued)

Box 10.4 *(continued)*

judicial procedures required to implement the dismissals. EPL usually establishes that either economic or disciplinary reasons have to be provided for dismissal by the employer, who has the burden of proof. Layoffs are considered unfair in most countries when there are neither subjective (misconduct) nor objective (economic) grounds for the interruption of the relationship. Penalties applied to employers implementing unfair dismissals do not discriminate among the two types of justifications (disciplinary and economic), and an employer who finds it hard to prove misconduct can always try to justify the dismissal on economic grounds. Thus the costs of disciplinary layoffs are unavoidably interrelated with those of economic dismissals. Overall, insofar as EPL negatively affects disciplinary layoffs, it increases the efficiency wage; when EPL instead acts mainly on economic layoffs, it reduces the efficiency wage. This suggests that the impact of EPL will vary depending on the monitoring technology of the firm. The better this technology, the stronger the effects of EPL on disciplinary versus economic layoffs. As small firms can better monitor the productivity of their workers than large firms, this provides an argument for the exemption of small units from the strictest EPL provisions (Boeri and Jimeno 2005).

Sources: Shapiro and Stiglitz (1984); Boeri and Jimeno (2005).

Other important design features of EPL concern, more than the level of protection, its unpredictability and the involvement of third parties. As discussed in this chapter, EPL consisting of a tax is always more distortionary than EPL designed as a pure transfer. However, some payments to third parties are unavoidable because judges need to be involved in assessing the nature (economic, disciplinary, or discriminatory) of a dismissal. Because the involvement of jurisprudence is unavoidable, the costs of layoffs for employers and workers are highly uncertain. One way to reduce uncertainty about court rulings is to put some limits on the arbitrariness and the duration, and hence the costs, of judicial involvement. For instance, the Dutch dual system allows workers involved in layoffs to benefit from a quicker, more predictable, and higher severance payment. An increasing proportion of employers (from less than 10 percent in the late 1980s to more than 50 percent more recently) indeed prefer to opt for a quicker but more costly dismissal procedure than to wait for the authorization of a designated public authority (the Centre for Work and Income), which does not require any severance payment in case of a fair dismissal.

10.4.2 Whom Should EPL Be Protecting?

A question of the International Social Survey Programme (ISSP), a household survey carried out in the EU-15 since 1994, is particularly valuable in highlighting

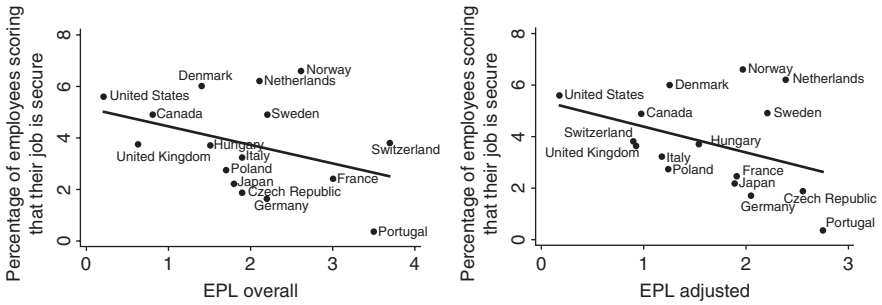


Figure 10.4 Perceptions of Job Security and Strictness of EPL

the coverage and extent of the protection offered by EPL. Employees with a permanent contract were asked how satisfied they were with their present job or business in terms of job security. As pointed out by Postel-Vinay and Clark (2005), the survey collected information on the respondent's perception of at least two different components of job security, (1) the probability of job loss and (2) the cost of job loss. Comparisons of the age and occupational profile of the answers suggest that the first factor dominates: it is indeed the persons less protected against the risk of job loss (e.g., workers in small firms or youngsters) who report lower degrees of satisfaction with their present job.

With this caveat in mind, the left panel of figure 10.4 measures, on the vertical axis, the fraction of employees in the 2005 ISSP stating that their job is secure in the various countries and, on the horizontal axis, the OECD index of strictness of EPL. Surprisingly enough, countries with the strictest EPL provisions (at least according to the OECD index, unadjusted by coverage), notably the southern European countries, have the largest fraction of workers who are deeply concerned about their job security. A possible interpretation of this result is that EPL is very selective in offering protection against the risk of job loss: this protection can only be offered to a limited subsegment of the workforce, concentrating the risk on those who are not covered. In other words, the high perceived risk may be a by-product of the fact that EPL protects only workers with permanent contracts and concentrates all risks on the others as a result of the dual-track reform strategies characterized in this chapter. However, even when we use the adjusted measure (by coverage) provided in table 10.2, the correlation between perceived security and EPL is negative (see figure 10.4, right panel). Another interpretation is that although individuals perceive a lower probability of job loss, they are aware that a job loss is more costly under strict EPL regimes because it involves longer unemployment spells. Yet another rather imaginative interpretation recently put forward by Wasmer (2006) is that under strict EPL employers use mobbing as a strategy to force their employees to leave the firm "voluntarily."

10.5 Why Does Employment Protection Legislation Exist?

Employment protection legislation is an institution that protects a limited segment of the workforce against unemployment risk. Because of the selectivity of the coverage offered by EPL, it is always better to complement measures of the strictness of regulations with information on the actual coverage of these norms. In some countries very strict EPL for regular contracts goes hand in hand with a large share of employment in the informal sector, in plants exempted from these norms, or in self-employment, while many workers are hired under temporary contracts. This may help explain why in countries with strict EPL, many workers do not feel at all less secure about their jobs.

For all these reasons, EPL is a strongly redistributive institution. It protects those who already have a job, notably a permanent contract in the formal sector. Unemployed individuals and workers with temporary contracts generally suffer in the presence of strict EPL rules for permanent contracts. The former tend to experience longer unemployment spells in the presence of strict EPL, while the latter may be caught in a sort of parallel secondary labor market of temporary contracts with limited access to the primary labor market of permanent contracts. Employers also suffer a loss in profits in the presence of employment protection, notably when they do not succeed in making workers pay (through lower wages) for the costs of providing this insurance.

From a political economy perspective, then, depending on the relative number and/or political power of (1) employees with permanent contracts, (2) the unemployed, (3) workers with temporary contracts, and (4) employers, we should have more or less stringent EPL. Another important factor affecting political support for EPL is the presence of other institutions, such as UBs, that can substitute for EPL in providing insurance against unemployment risk, as further discussed in chapter 13. Under stronger competition in product markets, a case can be made for protections against job loss that are more mobility friendly. UBs are one of these. Thus we should expect EPL to become less important over time as countries improve the coverage of their unemployment benefit systems.

10.6 Suggestions for Further Reading

The chapters on EPL in OECD (1999, 2004b) provide a good overview of the problems related to the measurement of this institution. Enforcement (and policy endogeneity) issues are discussed in some detail in Giuseppe Bertola, Tito Boeri, and Sandrine Cazes (2000). Tito Boeri, and Pietro Garibaldi (2007) analyze the honeymoon effect and offer a survey of the main findings of the literature on dual labor market regimes. Finally, Gilles Saint-Paul (1993) is a compulsory reference on the political economy of employment security.

10.7 Review Questions

1. What are the main drawbacks of available measures of the strictness of EPL?
2. Why is there a nonmonotonic relationship between EPL and unemployment?
3. Why many two-tier reforms of EPL transitorily increase employment?
4. What are the efficiency arguments in favor of employment protection?
5. When is EPL neutral with respect to labor market outcomes?
6. Why from an economic standpoint is it important to disentangle the transfer from the tax component of EPL?
7. Why, in your view, has empirical work failed to observe the negative relationship between EPL and job and labor turnover predicted by economic theory?
8. Why do workers in countries with strict EPL feel less secure than workers in flexible labor markets?
9. Why are third parties (e.g., judges) always involved in the enforcement of EPL?

10.8 Technical Annex: EPL in a Search-Matching Model

10.8.1 A Simple Dynamic Framework

A search-matching model originally developed by Boone and van Ours (2006) can be valuable in characterizing equilibriums in dynamic labor markets, as well as the role of labor market institutions. In a search-matching model unemployment arises because workers need time to find vacancies and employers need time to find unemployed workers to fill their vacancies. Labor is assumed to be homogeneous, and all jobs offer the same wage. If an unemployed worker and a vacancy are matched, they generate a constant productivity as long as the job is not destroyed by an exogenous shock. Individuals and firms are assumed to be risk neutral. The main elements of the baseline model are workers' behavior, matching of unemployed workers and vacancies, behavior of firms, wage bargaining, and labor market equilibrium flows.

10.8.2 Present Discounted and Flow Values

"I want it all and I want it now" are famous lyrics.¹ These lyrics represent an economist's way of thinking: maximize, and maximize preferably as soon as possible—better today than tomorrow. The *discount rate* is an indicator of this

¹ From Queen, "I Want It All," *The Miracle* (Burbank, CA: Hollywood Records, 1992).

line of thinking. It is the time rate at which economic value depreciates. The net present value of a continuous flow of wages w is

$$V_w = \int_0^{\infty} w e^{-\rho t} dt = \frac{w}{\rho}, \quad (10.5)$$

where ρ is the discount rate. In other words, the flow value of having a wage forever equals the product of the discount rate and the net present value:

$$w = \rho V_w. \quad (10.6)$$

10.8.3 Behavior of Workers

All workers are assumed to be homogeneous so that they receive the same wage w net of taxes. Unemployed workers search at a cost (disutility), γ .

The flow value of being unemployed equals the difference between the cost of search and the expected benefits from finding a job:

$$\rho V_u = -\gamma + \mu (V_e - V_u), \quad (10.7)$$

where μ is the job-finding rate. The flow value of having a job is

$$\rho V_e = w + \delta (V_u - V_e), \quad (10.8)$$

where δ is the exogenous job separation rate.

10.8.4 Matching

Vacancies and unemployed workers are matched by a *matching function* that describes the efficiency of labor markets. The matching function, for any given flow of matches, is equivalent to the so-called Beveridge curve:

$$m(u, v) = A(u)^{1-\eta} v^{\eta}. \quad (10.9)$$

In numerical simulations it is customary to specify the matching process as a Cobb-Douglas function with constant returns to scale and A and η as relevant parameters. The matching function can then be simply rewritten in terms of the number of vacancies divided by the number of unemployed workers (market tightness):

$$\mu = \frac{m(u, v)}{u} = A\theta^{\eta}, \quad (10.10)$$

where θ denotes labor market tightness,

$$\theta = \frac{v}{u}. \quad (10.11)$$

10.8.5 Firms

Given that production y is sold at (numeraire) price 1, the flow value of a job is

$$\rho J_e = y - w + \delta(J_v - J_e). \quad (10.12)$$

The flow value of a vacancy is

$$\rho J_v = -c_v + \frac{\mu}{\theta}(J_e - J_v), \quad (10.13)$$

where $\frac{\mu}{\theta}$ is the rate at which vacancies are filled and c_v is the cost of an unfilled vacancy.

10.8.6 Wage Bargaining

To determine the wage, the Nash-bargaining solution (see technical annex 3.8) is used. Notice that here we are dealing with individual workers rather than a union. The bargaining power of the worker is denoted by $\beta \in [0,1]$ and the bargaining power of the employer by $(1 - \beta)$. Then the net wage w solves

$$\max_w (V_e - V_u)^\beta (J_e - J_v)^{1-\beta}, \quad (10.14)$$

or

$$\max_w \left(\frac{w}{\rho + \delta} + \frac{\delta}{\rho + \delta} V_u - V_u \right)^\beta \left(\frac{y - w}{\rho + \delta} + \frac{\delta}{\rho + \delta} J_v - J_v \right)^{(1-\beta)}. \quad (10.15)$$

The first-order condition for w can now be written as

$$(1 - \beta)(w - \rho V_u) = \beta(y - w), \quad (10.16)$$

or, equivalently,

$$w = (1 - \beta)\rho V_u + \beta y. \quad (10.17)$$

The higher the worker's bargaining power (the higher the β), the higher the share of the surplus y . In case workers have full bargaining power ($\beta = 1$), the wage is equal to the total surplus. In case employers have full bargaining power ($\beta = 0$), the wage is equal to the flow value of unemployment ($w = \rho V_u$). In the first case the worker gets the entire surplus. In the second case the employer extracts all the surplus, conditional on the worker having some incentive to work.

10.8.7 Introducing EPL

We introduce EPL through firing costs c_f that are assumed to be a pure tax or deadweight cost; that is, they impose costs on the employer in case of firing but

are not paid to the worker through severance payments. It can readily be shown that a pure transfer under our assumptions (risk-neutral workers) cancels out in wage bargaining. An EPL tax instead affects wage bargaining,

$$w = (1 - \beta)\rho V_u + \beta(y + c_f), \quad (10.18)$$

vacancy creation,

$$J_e = \theta \frac{c_v + c_f}{\mu}, \quad (10.19)$$

and the value of a job,

$$J_e = \frac{y - w + c_f}{\rho + \delta}. \quad (10.20)$$

The most intuitive channel via which EPL affects the equilibrium is the job destruction rate. To keep things simple, we will assume that the job destruction rate linearly decreases with employment protection:

$$\delta = \delta_0 - \delta_1 c_f. \quad (10.21)$$

10.8.8 Baseline Parameters

The model for simulation exercises (some results are displayed in figure 10.3) has 11 endogenous variables, γ , V_u , V_e , w , J_e , J_v , v , u , m , θ , and μ , and 8 parameters for which the following baseline values are used: discount rate $\rho = 0.025$, matching function $\eta = 0.5$, $A = 1.0$, wage negotiations $\beta = 0.5$, match productivity $y = 1$, vacancy costs $c_v = 2$, job destruction rate $\delta_0 = 0.04$, and $\delta_1 = 0.1$. Finally, since there is free entry in posting vacancies, $J_v = 0$.

11

Unemployment Benefits

Unemployment benefits (UBs) protect individuals against uninsurable labor market risk. Unlike employment protection legislation (covered in chapter 10) that protects jobholders by making it more difficult for employers to dismiss their workers, UBs offer replacement income to workers experiencing unemployment spells after having lost their job.

The first unemployment benefit system was introduced in the United Kingdom in 1911. Many people did not like the idea that persons not working could receive a state transfer, and beneficiaries were ironically named people “on the dole.” Nowadays all OECD countries, most middle-income countries, and some developing nations have an unemployment benefit system in place.

Unemployment benefits are often treated as a one-dimensional institution. However, there are several key dimensions that identify an unemployment benefit system: *eligibility* (the norms determining access to the benefit) and *entitlement* (the rules concerning the duration of the payment) conditions and level of payments. Typically, job losers with short unemployment durations can have access to an income replacement system that mimics an insurance scheme: they draw benefits that are proportional (at least above a benefit floor and below a given threshold) to their past contributions. These contributions are generally channeled to an extrabudgetary *unemployment insurance fund* that collects compulsory payroll taxes earmarked to unemployment benefits. This *unemployment insurance* (UI) component of UBs provides transfers that are proportional to the last wage earned (hence to past contributions) and whose duration increases with the length of the contribution record. The close link between benefits and past contributions, entitlements, and insurance premiums paid during the working life reduces incentives to elude or evade payments to the social insurance fund.

Transfers to jobseekers with longer unemployment durations are generally independent of past contributions and are offered in combination with other cash transfers to individuals who are not working, notably social assistance of the last resort. Eligibility to this second, *unemployment assistance* (UA) component of UBs can be independent of payments made during previous work experience. UA is typically offered at a flat rate (independent of the previous wage) and for a

maximum duration that is not conditional on the length of the contribution period. UA can also be integrated with general social assistance, in which case it is offered for unlimited duration, but subject to means testing; that is, it is provided only to unemployed individuals who have incomes and family assets lower than a given (poverty) threshold. Linking the transfer to household assets involves tying together individual and family labor supply decisions. For this reason, we discuss the effects of means testing on family labor supply elsewhere (chapter 4).

UBs affect labor force participation, employment, and unemployment via a variety of channels that involve interactions with other institutions such as collective bargaining institutions (chapter 3) and active labor market programs (chapter 12). Because payroll taxes have to be levied in order to finance the benefits, the effects of UBs can be better understood by considering the effects on the labor market equilibrium of taxes on labor (chapter 4).

11.1 Measures and Cross-Country Comparisons

Mapping all the various features of UBs into a one-dimensional measure is not an easy task. Table 11.1 tabulates the *net replacement rate*, that is, the ratio of the unemployment benefit to previous earnings at different earning levels, both measured after taxes. Replacement rates net of taxes tend to be higher than gross replacement rates because income taxes are progressive and UBs in some countries are tax exempt. Because of the progressivity of taxation and the presence of benefit floors and ceilings, replacement rates of high-wage earners (earning 150 percent of the wage of the average production worker, APW) tend to be significantly lower than the replacement rates offered to those coming from the low end of the wage distribution (those receiving two-thirds of the average wage, as shown by the first two columns on the left-hand side of the table). Another important design feature of UBs highlighted by table 11.1 is the fact that replacement rates are generally lower at the 60th month of unemployment than at the beginning of the unemployment spell. This design feature, which is more evident for those earning at least the APW wage, corresponds to optimal unemployment insurance principles requiring that replacement rates decline over the unemployment spell in order to provide incentives to search for a job at short unemployment durations. Relatively high replacement rates in the first months of unemployment allow individuals to concentrate immediately after the layoff on job search without having to drastically modify their living standards (e.g., sell their house or their car). As long as the level of the benefit is declining over time, the worker will be encouraged to actively search for a job before that worker's human capital depreciates.

There are, however, a few cases highlighted by table 11.1 (e.g., Austria, Germany, and Ireland) in which the replacement rate increases over time for low-income earners. This happens because long-term unemployed workers become

eligible for unemployment assistance or social assistance that can be more generous than UI for those coming from the low end of the wage distribution (50 percent of two-thirds of the APW wage can well be below the poverty line).

As table 11.1. suggests, it is difficult to make cross-country comparisons of replacement rates because these vary considerably across earning levels and at different unemployment durations. What replacement rate should be taken as the

Table 11.1 Net Replacement Rates at Different Income Levels and Unemployment Durations

	67% of APW		100% of APW		150% of APW	
	First month (1)	60th month (2)	First month (1)	60th month (2)	First month (1)	60th month (2)
Australia	46	46	32	32	24	24
Austria	55	64	55	51	55	51
Belgium	87	72	66	55	49	40
Canada	63	31	64	22	44	15
Czech Republic	50	45	50	31	50	21
Denmark	84	71	59	50	45	37
Finland	78	69	64	51	50	37
France	80	56	71	41	70	29
Germany	63	82	61	61	62	54
Greece	64	0	46	0	33	0
Hungary	61	33	44	24	34	18
Iceland	66	66	49	49	35	35
Ireland	40	71	29	51	22	38
Italy	50	0	52	0	46	0
Japan	73	50	63	34	62	23
Korea	53	25	54	17	42	12
Luxembourg	84	70	85	50	87	37
Netherlands	79	79	71	58	61	39
New Zealand	54	54	37	37	26	26
Norway	66	60	66	42	53	31
Poland	65	45	44	30	30	21
Portugal	85	34	78	24	83	17
Slovak Republic	69	62	62	42	44	29
Spain	76	37	70	27	48	19
Sweden	82	74	81	51	58	37
Switzerland	79	73	72	51	72	36
United Kingdom	63	63	45	45	31	31
United States	60	10	56	7	40	5

Source: OECD, *Benefits and Wages*, various issues.

Note: The net replacement rates are estimated by taking benefits and wages after taxes and including in the benefits also social assistance and housing benefits offered to unemployed individuals.

reference in international comparisons of UB generosity? The OECD tabulates a *summary measure of benefit generosity*, which is defined as the average of the replacement rates in the first four years of unemployment for an APW (table 11.2, second column). This cutoff duration (48 months) is clearly arbitrary. It takes into account that the average length of an unemployment spell is lower than two years in most OECD countries and that after four years jobless people lose their attachment to the world of work. This summary measure clearly attaches the same importance to replacement rates offered at different unemployment durations and ignores the time profile of replacement rates. To give an example, a system offering 50 percent of the previous wage throughout the entire unemployment spell would be considered as generous as one offering 65 percent in the first year, 55 percent in the second, 45 percent in the third, and 35 percent in the fourth year of unemployment.

A more serious problem with the OECD summary measure is that it neglects *eligibility* (the norms determining access to the benefit) and *entitlement* (the rules concerning the duration of the payment) conditions. Because of strict eligibility conditions and short entitlement relative to the typical duration of unemployment spells, the coverage of unemployment benefits (the fraction of unemployed workers receiving the benefits) can sometimes be very low. Unfortunately, it is difficult to obtain cross-country-comparable measures of coverage. For the EU-15 countries, UB coverage can be computed by drawing on the European Community Household Panel (ECHP, which was regrettably discontinued in 2001). Table 11.2 presents the OECD summary measure of generosity (based on legal norms on replacement rates) in the first column, while in the second column it displays the fraction of individuals who are unemployed according to OECD-ILO definitions (see box 1.3) and who declared that they were beneficiaries of UBs in the 2001 ECHP. An adjusted summary measure is then displayed in the third column that corrects the OECD measure by coverage. This adjusted measure is obtained by multiplying the OECD summary generosity measure by the coverage ratio. Table 11.2 suggests that countries like Italy and Spain offer a much less generous UB system when account is taken not only of replacement rates, but also of eligibility and entitlement conditions.

The preceding discussion suggests that UB generosity measures unadjusted by coverage can be fairly misleading. At the same time, one should keep in mind that legal replacement rates and coverage rates of UBs are not independent of one another. Coverage is endogenous to replacement rates via take-up incentives (the more generous the UB system, the higher the benefit take-up) and fiscal constraints (the more generous the statutory replacement rate, the stronger the incentive of governments to tighten eligibility criteria, reducing the coverage of UBs). The trouble is that the first type of correlation is positive (more generous benefits encourage more people to apply for UBs) while the second source of correlation, combining fiscal and political economy constraints, induces a negative

Table 11.2 Adjusting the OECD Generosity Measure by UB Coverage

Country	Unadjusted OECD generosity measure (a)	Coverage of UBs ¹ (b)(%)	Adjusted, (a)*(b)
Austria	0.82	31.5	25.8
Belgium	0.82	38.5	31.5
Denmark	0.85	50.9	43.3
Finland	0.82	34.8	28.5
France	0.74	43.5	32.0
Germany	0.83	29.6	24.7
Greece	0.38	13.0	4.9
Ireland	0.88	25.8	22.8
Italy	0.18	34.1	6.1
Netherlands	0.54	52.9	28.6
Portugal	0.39	41.2	16.2
Spain	0.48	36.5	17.6

Sources: OECD, *Benefits and Wages*, various issues, for the data on net replacement rates. Our estimates of the coverage rate are based on the European Community Household Panel.

¹Coverage is the fraction of LFS unemployed declaring that they were receiving unemployment benefits. The fourth column is the product of the numbers in the second and third columns.

correlation between replacement rates and benefit coverage. Governments may pursue a policy of low coverage with very generous soft-landing schemes offered to a few job losers, leaving many other displaced workers, typically those less politically represented, with low, if any, income support. This combination seems to prevail in many middle-income countries. UBs there offer relatively high nominal replacement rates (e.g., 60 percent of the best earnings in the last year in Argentina), but are offered for a short period of time and only to a small fraction of the workforce (workers in small businesses and in rural areas are not covered, as in China). These asymmetries in replacement rates and duration of benefits are somewhat less evident, but still present, in OECD countries. In southern Europe, for instance, UBs are relatively generous in terms of replacement rates but cover a very small fraction of the unemployed, as documented by table 11.2. Conversely, in the United Kingdom and even more in the United States, low replacement rates go hand in hand with almost universal coverage of job losers and are generally offered—when integrated with means-tested social assistance—with unlimited duration.

The preceding discussion offers strong arguments against the standard practice of mapping unemployment benefits into a simple scalar generosity measure. It is always better to use a variety of measures to characterize UBs. Since they are multidimensional institutions, UBs can be reformed along various dimensions. The devil is in the details, and neglecting any of these dimensions can offer a misleading picture of the relative generosity (and evolution over time) of a UB system.

11.2 Theory

11.2.1 A Competitive Labor Market

Most textbooks discuss the effects of UBs only in competitive labor markets. As is argued here, this discussion is not particularly instructive because unemployment (and hence unemployment benefit systems) does not arise as an equilibrium phenomenon in this context. Yet it may be useful to start from the competitive market insofar as it helps in understanding some design features of UBs that can mitigate its adverse effects in more realistic environments.

In competitive labor markets unemployment benefits affect labor market outcomes along two main channels:

1. They increase the (static) reservation wage of individuals, inducing more people not to supply labor at any given market wage (the *participation effect*).
2. They affect labor market outcomes via the increased taxation (generally payroll taxes) required to finance UBs (the *taxation effect*).

The second transmission mechanism is not discussed in this chapter because it was already investigated in detail in chapter 4. Yet it is important to keep in mind that any increase in the generosity of UBs needs to be financed, and that UBs are generally funded via statutory payroll contributions. This means that more generous UBs generally involve an increase in *proportional* taxation of earnings. When UBs are funded via taxation of incomes, UB systems are more progressive.

In order to characterize the participation effect, we need to refer to the notion of the *reservation wage*, which was introduced in chapter 1. Each individual has a utility function defined over consumption, c , and leisure, l , which are both assumed to be normal goods. The utility function is concave in c , which means that individuals are risk averse. Thus their welfare can be increased by giving them the opportunity to purchase some (actuarially) fair insurance scheme that reduces their uncertainty about the income/leisure pairs to which they can have access. This is indeed the insurance (or income-smoothing) role attributed to UBs.

In order to keep things simple, however, we shall consider an environment in which there is no uncertainty as to potential incomes and labor/leisure levels. The individual then allocates the endowment of time, say, l_0 , alternatively to work, earning the hourly wage w times the number of hours of work, h , or to leisure (clearly $h = l_0 - l$). Define the individual's nonlabor income (the income when working zero hours) as m . As long as m is strictly positive, there will be a kink in the budget constraint of the individual: at zero hours of work (when $l = l_0$), the individual gets m . Income increases to the left of l_0 (for any positive amount of work) at the rate w , as depicted in the left-hand panel of figure 11.1. The individual maximizes welfare by choosing the highest possible indifference curve

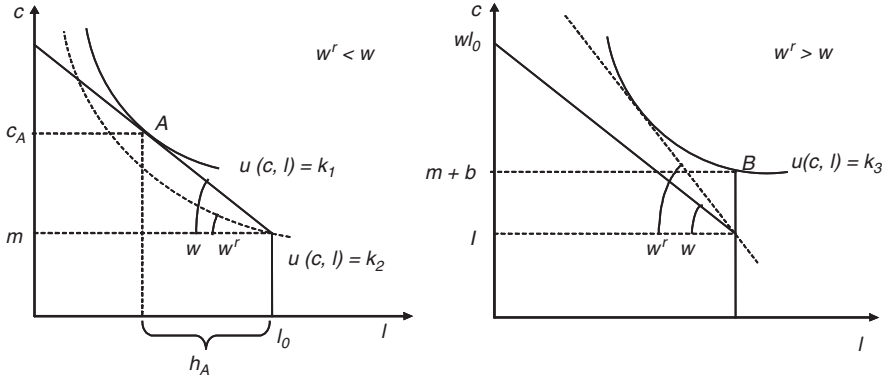


Figure 11.1 Reservation Wage without (Left Panel) and with (Right Panel) Unemployment Benefits

that is in the budget constraint, that is, the locus A where the indifference curve $U(c, l) = k_1$ is tangent to the budget constraint. The (static) reservation wage, w^r , is given by the slope of the indifference curve $U(c, l) = k_2$ crossing the kink of the budget constraint, where the individual allocates nonlabor income m to the purchase of consumption goods and works zero hours, as formally characterized in box 11.1.

Box 11.1 Reservation Wages and Unemployment Benefits

As also discussed in chapter 1, the reservation wage is formally defined as the *marginal rate of substitution* between leisure and consumption at the kink of the budget constraint, or

$$\frac{U_l(m, l_0)}{U_c(m, l_0)} = w^r.$$

Notice that, in general, the slope of the indifference curve at the kink will differ from the slope of the budget constraint, which is given by the market wage, w . Now, for any $w > w^r$, that is, any market wage exceeding the reservation wage, the individual will supply some positive hours of work. This is the scenario depicted in the left panel of figure 11.1, where the individual maximizes welfare at A , where $h_A > 0$ hours of work are supplied. If instead $w < w^r$, then the worker will not supply any positive hour of work because that worker can be better off devoting time entirely to leisure and consuming m .

It may be more realistic to consider a case where there is a constraint in the total number of hours to be worked, for example, that the individual deciding to supply labor is forced to work h_{ft} hours, corresponding to a full-time

(continued)

Box 11.1 (continued)

position. As discussed also in chapter 1, the reservation wage will now implicitly solve the condition

$$U(m, l_0) = U(m + w^r h_{ft}, l_0 - h_{ft}).$$

Suppose now that an unemployment benefit, b , is introduced. Nonlabor income will now be $m + b$. Notice that the benefit is conditional on not working. Thus, in the presence of unemployment benefits, the reservation wage is implicitly given by

$$U(m + b, l_0) = U(m + w^r h_{ft}, l_0 - h_{ft}). \quad (11.1)$$

Since utility is strictly increasing in both arguments, the reservation wage will be increasing with the level of unemployment benefits: the larger the b , the stronger the increase of w^r after the introduction of the UB system. This negative participation effect can be partly mitigated by allowing the individual to combine the unemployment benefit with work income, at least partly, making the benefit also conditional on employment at low wages, as discussed in chapter 12.

Suppose now that an unemployment benefit is introduced providing some positive income, b , to those who are not working in addition to m and that this additional income is provided conditional on not working. The new budget constraint will now look as in the right panel of figure 11.1: it will have a spike at the zero hours locus because the transfer b is provided only to those working zero hours. Provided that the transfer is sufficiently large, the individual may now be better off not working at all. Indeed, in the right panel of figure 11.1, the individual can now attain a higher indifference curve ($k_3 > k_1$) than at the equilibrium without UBs. When UBs are present, the reservation wage is no longer represented by the slope of the indifference curve at the kink. It will now be given instead by the lowest market wage that can allow the individual to reach at least the indifference curve $U(c, l) = k_3$ by supplying some positive hours of work. Because now $w < w^r$, the marginal value of leisure exceeds its opportunity cost (the market wage), and hence the individual is better off not working. In order for the individual to supply labor, hourly wages should therefore increase. Moreover, the assumption that leisure is a normal good implies that any increase in nonlabor income, for example, an increase brought about by a higher unemployment benefit, increases the reservation wage.

All in all, the prediction of a simple static neoclassical model of labor supply is that the introduction of an unemployment benefit system increases the reservation wage of individuals and thus the value of the nonparticipation option. When leisure is a normal good, the larger the unemployment benefit, the stronger the increase in the reservation wage with respect to the case where no UB is present, and hence the more likely that the UB recipient will decide not to work.

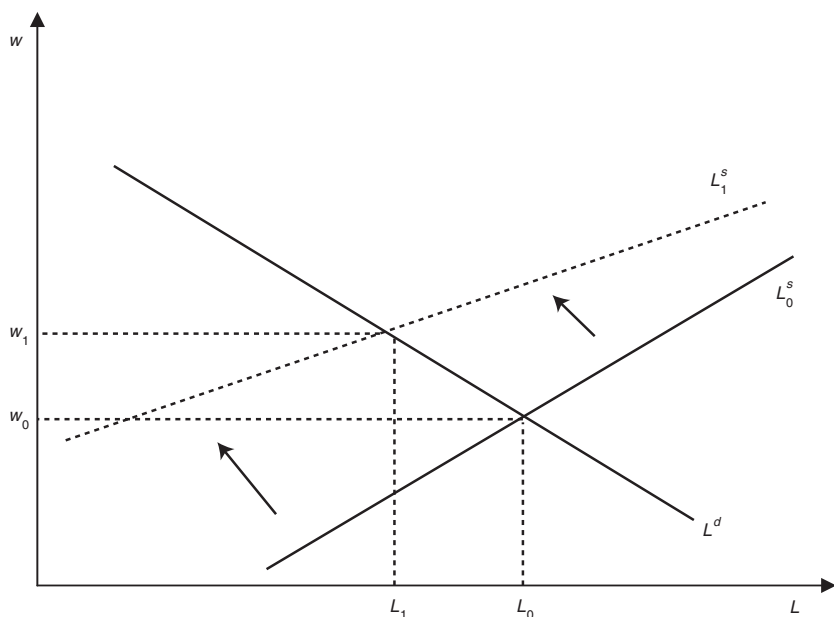


Figure 11.2 Aggregate Labor Supply with (Dotted Line) and without (Continuous Line) Unemployment Benefits

What does the increase in the reservation wage imply for aggregate labor market outcomes? As long as workers have different preferences over the labor/leisure choice, there will be different reservation wages in this economy. This means that labor supply will be increasing with the market wage, L^s , as in figure 11.2 and in the head-count labor supply derived empirically in chapter 1 from the cumulative density function of the stated reservation wages of individuals. The higher the market wage, the larger the fraction of the population of working age willing to supply labor.

Suppose now that an unemployment benefit is introduced. We know from the preceding discussion that this will increase all reservation wages, shifting labor supply upward, L_1^s as in the dotted line of figure 11.2. It is likely that the effect is stronger for individuals with a lower reservation wage; hence the distance between the labor supply with and without UBs will be larger the closer we are to the origin. It follows that the labor market equilibrium with unemployment benefits involves a higher wage ($w_1 > w_0$) and lower employment ($L_1 < L_0$) than the (wedge-free) equilibrium without unemployment benefits.

Notice that in this simple framework unemployment benefits reduce the size of the labor market, but do not give rise to unemployment. In other words, UBs act like an increase in compulsory schooling, reducing labor supply. The fact that

unemployment does not arise at the equilibrium makes the competitive model not particularly interesting in analyzing the effects of UBs. If unemployment is simply not there, then why should there be an unemployment benefit system?

11.2.2 An Imperfect Labor Market

In labor markets with some imperfections, such as frictions in the job creation process, wage bargaining, or imperfect information, unemployment benefits affect labor market outcomes via three main channels:

1. They increase the reservation wage of unemployment benefit recipients because they make workers more choosy when deciding upon available job offers, reducing their search intensity and hence increasing the duration of their unemployment spells (the *job search effect*).
2. They improve the fallback option of workers, putting a higher floor in wage setting and increasing wage claims at the bargaining table, or in any event the wages required to deter shirking (the *wage effect*).
3. They induce more people to participate in the labor market insofar as they increase the value of employment and unemployment over the value of inactivity (the *entitlement effect*), at least for those not receiving the benefit.

In addition to these effects there will be also the *taxation effect* that is present also in competitive labor markets. These effects can be characterized in a dynamic framework as the model formally described in technical annex 11.8. We confine ourselves here to providing the intuition behind the operation of these three transmission mechanisms.

11.2.2.1 The Job Search Effect

The static model of labor supply, where an individual optimally chooses hours of work and leisure subject to a budget constraint, is ill suited for an analysis of the effects of UBs on unemployment duration because it does not allow unemployment to arise as an equilibrium phenomenon. In order to fully characterize the effects of UBs, we need a more realistic, but also somewhat more complex, model. *Job search theory* (pioneered by Lippman and McCall 1979) provides such a framework. It pays attention to the time and costs involved in looking for jobs. This process is modeled like the search for an apartment: one surveys the market and does not necessarily accept the first offer, but keeps on searching until one finds something that corresponds to one's expectations. Bringing the theory closer to the labor force statistics, these models also allow individuals to decide upon the amount of effort to put into job search, for example, in terms of hours of leisure sacrificed to the search for a good job. A person who decides not to search at all

is inactive. Those actively seeking a job are unemployed. It is also possible to build a measure of *effective labor supply* that goes beyond simple head counts of jobseekers and considers the intensity of job search of different individuals. Under some standard assumptions (jobseekers form expectations about the wage offers that they are likely to receive because they know the distribution of wages from which job offers are randomly drawn, each job pays a constant wage, and there is no on-the-job search so that one cannot accept a job and continue to search), it turns out that the optimal search strategy involves a *stopping rule*, that is, defining a threshold wage level such that only the jobs that offer a wage superior to this level are accepted. This threshold wage level is the (dynamic) *reservation wage* of the individual and equals the value of unemployment because it makes the individual indifferent between working at that wage and continuing the job search. It is clearly a different concept from the (static) reservation wage characterized in box 11.1. The static reservation wage discriminates between employment and non-participation, while in job search theory one continues to participate in the labor market even if the wage offer is lower than the reservation wage. One just keeps on searching. Thus, the dynamic reservation wage of job search theory discriminates between employment and unemployment.

The model developed in technical annex 11.8 involves a *two-sided job search* in that not only workers but also firms are searching. Employers issue vacancies that will be gradually filled or matched with jobseekers. Employers' searches are also costly. Moreover, jobs do not last forever, but are destroyed at some given rate. The model generates labor market flows across employment and unemployment whenever a jobseeker is matched to a vacancy or a job is destroyed. Equilibrium involves some positive level of unemployment. Actually, some unemployment is efficient because it makes it easier for employers to fill their vacancies. There are *externalities* involved in the matching of vacancies and jobseekers. The more vacancies around, the easier it is to find a job. Similarly, the more unemployed workers seeking a job, the easier it is to fill a vacancy.

While searching for jobs, individuals earn an unemployment benefit that is proportional to the wage, as in earnings-related UB systems.¹ In this framework an increase in the generosity of unemployment benefits unambiguously reduces the optimal level of search intensity. This happens because the opportunity cost of having a job increases. As a result of the decline in search intensity, the job-finding probability declines as well, increasing the duration of unemployment.

These negative effects of unemployment benefits on job search can be reduced by offering benefits only for a limited period of time. This implies that the negative effects on job search intensity decrease together with the length of the spell of joblessness, mitigating the effects of unemployment benefits on unemployment

¹ This assumption is justified by the fact that individuals are homogeneous and all earn the same wage in the model in technical annex 11.8.

duration. This possibility is discussed in chapter 12, where we analyse active labor market policies that set de facto a limit on the maximum duration of benefits.

11.2.2.2 Wage Effects

Because there are costs in job search and in issuing vacancies and the matching of jobseekers and vacancies is time consuming, even if product markets are competitive (and hence the expected profits from posting a vacancy are zero at the equilibrium), filled vacancies generate some positive rent for the employer and the worker involved. These rents are split between the employer and the worker. In particular, wages (establishing the fraction of the surplus going to the worker) are set according to a Nash-bargaining rule similar to that considered in chapter 3. Also in this case wages will be increasing with the fallback option of workers and hence with the unemployment income and the job-finding probability per any given level of search intensity.

Thus, in addition to reducing search intensity, unemployment benefits increase the equilibrium market wage via the Nash-bargaining rule. This effect will be larger the stronger the effect of UBs on the fallback option of workers, that is, the value of unemployment. As we have seen, the value of unemployment increases with the generosity of unemployment benefits because jobseekers get a higher income per any given level of search effort. However, more generous unemployment benefits reduce the vacancy rate because it is more difficult to fill vacancies when people search less intensively. The increase in wages itself reduces the job-finding probability per any given level of search intensity.

The wage effects of unemployment benefits are magnified by the taxes required to finance the benefits. Because benefits are funded via payroll taxes, the contributions clearing the unemployment budget should equal the unemployment benefit multiplied by the total number of unemployed individuals, divided by the total wage bill; that is,

$$t = \frac{bU}{wL}.$$

Unemployment benefits may increase the equilibrium market wage also because they force employers to pay their employees more in order to deter them from shirking, as in the efficiency wages literature, summarized in box 11.2.

Box 11.2 Efficiency Wages and Unemployment Benefits

Like job search theory, efficiency wage models (Shapiro and Stiglitz 1984) also generate involuntary unemployment as an equilibrium phenomenon as discussed in box 10.4. Employers who have only imperfect information about the effort of

(continued)

Box 11.2 (continued)

their employees pay a wage above the market-clearing level and use involuntary unemployment as a deterrent to shirking. As in the model in technical annex 11.8, the value of a job to a person who does not shirk is

$$\rho V_e^N = w - e + \delta(V_u - V_e^N), \quad (11.2)$$

where w is the wage, e is the effort, δ is the exogenous job separation rate (independent of the behavior of the employee), and V_u is the value of being unemployed. For shirkers, the flow value of a job is

$$\rho V_e^S = w + (\delta + q)(V_u - V_e^S), \quad (11.3)$$

where q is the rate at which shirkers are detected and fired. The efficiency wage makes the value of shirking (11.3) lower than or equal to the value of not shirking (11.2); that is,

$$w^* \geq \rho V_u + e \left(1 + \frac{\rho + \delta}{q} \right). \quad (11.4)$$

From this it follows that the efficiency wage is higher if the flow value of unemployment is higher and hence it hurts less to lose a job. If we ignore search, the flow value of unemployment is given by

$$\rho V_u = b + \bar{\mu}(V_e - V_u), \quad (11.5)$$

which is increasing with b , the level of unemployment benefits, and the job-finding rate, $\bar{\mu}$. At the equilibrium, inflows into unemployment must equal outflows from unemployment; that is, $\bar{\mu}U = \delta L$. It follows that $\bar{\mu} + \delta = \delta/u$, where U represents the number of unemployed workers, L the number of employed workers, and u is the unemployment rate. Using this condition, we can rewrite the efficiency wage as

$$w^* \geq b + \left(\frac{\rho + \delta/u + q}{q} \right) e. \quad (11.6)$$

Thus the efficiency wage increases with the level of unemployment benefits and decreases with the unemployment rate. If the unemployment rate is higher, a lower efficiency wage is needed. This is the sense in which unemployment is a worker-discipline device. A higher unemployment rate makes losing a job more costly and therefore encourages workers not to shirk. Unemployment benefits have the opposite effect: they reduce the penalty associated with being detected shirking and fired.

Source: Shapiro and Stiglitz (1984).

11.2.2.3 The Entitlement Effect

Labor supply can also vary as a result of a change in unemployment benefit generosity. In addition to being employed or unemployed, individuals may decide

not to participate at all in the labor market. This choice also has a reservation property: the individual decides to be inactive if that individual's nonlabor income while not working and not searching is higher than the reservation wage.

When unemployment benefits are paid only to active jobseekers, that is, inactive people do not receive the benefit, an increase in the generosity of benefits increases the reservation wage while leaving the value of inactivity unchanged. This means that some individuals will move from inactivity to unemployment, increasing labor force participation in the aggregate. The intuition is that the entitlement to higher income while seeking jobs induces more people to be engaged in active job search. Thus UBs may actually increase participation, contrary to the predictions of the competitive labor market model. This does not necessarily mean that labor supply also increases, because jobseekers put less effort into job search when unemployment benefits are higher.

When unemployment benefits are offered only to individuals with previous work experience, there may also be an entitlement effect among the population of jobseekers, that is, not only among inactive persons. When eligibility for benefits requires previous work experience, there are two categories of unemployed: first-time jobseekers, who do not receive the benefit, and jobseekers with previous work experience, who are eligible for UBs. The former group will actually increase their search intensity after the introduction of UBs because the value of holding a job increases (because it also involves eligibility for benefits in case of job loss), while for the second category there will be the usual negative effect on job search.

11.3 Empirical Evidence

11.3.1 Effects on Reservation Wage

The effects of UBs on the reservation wage are estimated by comparing stated reservation wages of unemployment benefit recipients and jobseekers not receiving UBs while controlling for personal characteristics, such as age, gender, and education, along the lines of box 1.2 in chapter 1.

The reliability of stated reservation wages is often challenged on the grounds that it may be affected by so-called hypothetical bias, whereby the interviewee does not have a clear notion of the type of job being considered. The effects of UBs on the reservation wage can also be identified by using *structural models* that make inferences about the implicit reservation wage of individuals on the basis of their job acceptance behavior. The studies applying this method typically found a positive elasticity of the reservation wage with respect to the level of UBs. For instance, Lynch (1983) reports elasticities of 0.08 to 0.11 for young people in the United Kingdom, Holzer (1986) finds elasticities of 0.018 to 0.049 in the United States, and van den Berg (1990b) of 0.04 to 0.09 in the Netherlands.



Figure 11.3 Maximum Duration of UBs and Long-Term Unemployment in OECD Countries

11.3.2 Effects on Unemployment Duration: Aggregate Data

The simple search model sketched in the preceding discussion and formally described in the technical annex predicts that UBs unambiguously increase unemployment duration. Consistent with this prediction, the maximum duration of UBs and the average duration of unemployment are generally found to be positively correlated across countries (see figure 11.3, drawn from Boeri, Layard, and Nickell 2001).

Several empirical studies drawing on aggregate data also looked at the relationship between generosity of UBs and level of unemployment. As stressed earlier, we would generally expect unemployment to rise in the presence of more generous UBs as a result of lower outflows from unemployment. This effect can be mitigated when benefits are paid only for a definite period of time and eligibility is conditional on past employment, in which case UBs may actually reduce unemployment duration among first-time jobseekers via the entitlement effect.

Using cross-sectional data on 20 OECD countries, Layard, Nickell, and Jackman (1991) found that countries with higher replacement rates and longer benefit duration tend to experience higher unemployment rates. In particular, they estimated that a 10 percent increase in the replacement rate involves roughly a 1.7 percent increase in the unemployment rate. More recent studies of the same group

of industrialized countries offer comparable results: Scarpetta (1996) estimated an elasticity of unemployment to UBs of the order of 1.3 percent, Nickell (1997) of 1.1 percent, and Bassanini and Duval (2006) of 1.2 percent.

In general, it is the maximum duration of benefits, as opposed to the level of the replacement rate, that has the strongest effect on unemployment rates. In other words, an increase of benefit levels has less effect on unemployment duration than an increase by the same percentage of the maximum duration of benefits. Unemployment benefits are also found to affect the composition of employment by pricing out women, youth, and older workers (Bertola, Blau, and Kahn 2002).

11.3.3 Effects on Unemployment Duration: Microeconomic Data

An advantage of microeconomic studies versus the macroeconomic literature is that they can disentangle the various transmission mechanisms associated with the introduction of UBs. They can also investigate the effects of UBs on different categories of jobseekers, notably those receiving UBs (whose reservation wage should increase together with benefits) and those not receiving them (whose reservation wage should decline together with UBs). Hence the effects on the average length of unemployment spells are likely to depend on the relative importance of these two groups of jobseekers.

A large empirical literature has been assessing the effects of benefit levels and residual entitlement on the duration of unemployment on the basis of longitudinal data. This literature typically applies *duration analysis* (Kiefer 1988), which extends to economics methods developed in industrial engineering (Lancaster 1979; Nickell 1979). The key measure analyzed by this literature is the *hazard* rate from unemployment, that is, the conditional probability of leaving unemployment in a particular week, given that one has been unemployed up to the beginning of that week.

This literature generally finds positive effects of UB generosity on unemployment duration, but the effects are quantitatively small, notably when UB generosity is measured in terms of replacement rates rather than in terms of the maximum duration of benefits.² To give a few examples, Narendranathan, Nickell, and Starn (1985) obtained a lower-bound estimate of only 0.08 for the effects on unemployment duration among British men of a 1 percent increase in the level of benefits. Van den Berg (1990) estimated that a 10 percent increase in the level of benefits in the Netherlands increases the duration of unemployment by one week, a result that is in line with evidence from the United States (Katz and Meyer 1990; and Mayer 1995b). The effect is, however, stronger (up to five weeks) when the increase in the benefit level occurs later in the unemployment spell.

² See Atkinson and Micklewright (1991), Devine and Kiefer (1991), and Krueger and Meyer (2002) for excellent surveys of this literature.

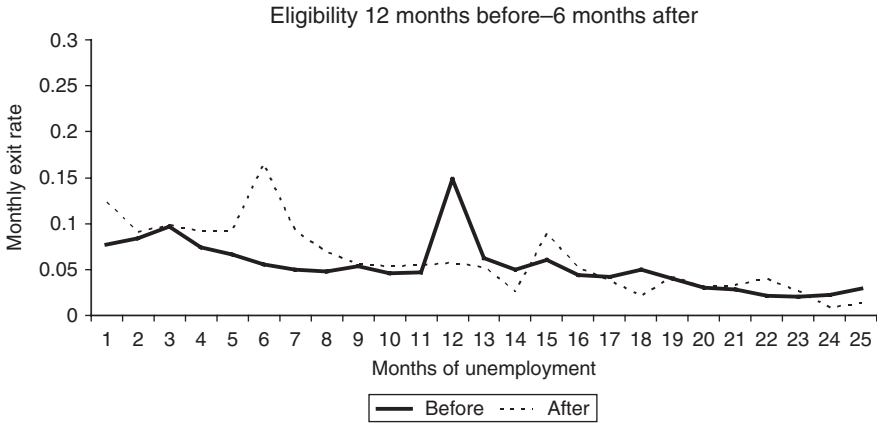


Figure 11.4 Hazards from Unemployment before and after the UB Reform in Slovenia (see box 11.3)

Stronger effects are found when one considers the effects of changes in the maximum duration of benefits (Meyer 1990). Several studies found spikes of exits from unemployment in correspondence with the maximum duration of benefits. Figure 11.4 illustrates these effects. It is drawn from Van Ours and Vodopivec (2006), who provided a difference-in-differences analysis of reforms of the UB system in Slovenia (see box 11.3). The figure suggests that reductions in the maximum duration of benefits (from 12 to 6 months) moved the spike in the hazard rate from the 12th to the 6th month. This result is in line with predictions of job search models when unemployment benefits are provided for a fixed duration and individuals can adjust search intensity (Mortensen 1977). Lalive, Van Ours, and Zweimuller (2005) also used difference-in-differences techniques to disentangle the effects on unemployment duration of increased replacement rates and the extension of the maximum duration of benefits. They found that the effects of benefit duration were stronger than the effect of the replacement rate (see box 11.4).

Box 11.3 Shortening the Duration of Benefits

Faced with increasing unemployment, Slovenia in October 1998 drastically reduced the potential UI benefit duration (PBD) roughly by half for most groups of recipients. Before reform, for example, workers with 15 to 20 years of experience were eligible for up to 18 months of benefits. After the change this group of workers was eligible for up to 9 months of benefits. Van Ours and Vodopivec (2006) analyze the effects of shortening the PBD. They identify a significant increase in the job-finding rate at various durations of unemployment spells. They also

(continued)

Box 11.3 *(continued)*

identify a clear spike in the job-finding rate in the month unemployment benefits expire. As shown in fig 11.4, the benefit spikes in the job-finding rates indeed follow the change in PBD. Whereas the probability of finding a job greatly increased for most groups of recipients whose benefit entitlement period was shortened, it stayed virtually unchanged for recipients whose entitlement period did not change. This suggests that the job-finding rate improved mainly because of more intense job search efforts by unemployed workers. In a follow-up study van Ours and Vodopivec (2008) find that the faster job finding did not occur at the cost of quality of postunemployment jobs; that is, workers found jobs faster without accepting lower-quality jobs in terms of stability, type of contract, and lower wages. Together, these findings imply that longer PBDs contributed to longer durations of unemployment spells of benefit recipients without improving the quality of postunemployment jobs. These findings thus imply that additional job search had a zero marginal productivity—or that recipients behaved opportunistically and did not spend additional time on job search at all, perhaps in collusion with their prospective employers.

Source: Van Ours and Vodopivec (2006).

Box 11.4 The Effects of Changes of Replacement Rates
and Maximum Benefit Duration in Austria

In 1989 Austria reformed its unemployment benefit systems in a way that affected various unemployed workers differently: one group experienced an increase in the replacement rate, a second group experienced an extension of the maximum duration of benefits, a third group experienced both a higher replacement rate and a longer duration of benefits, and a fourth group experienced no change in the policy parameters. In particular, unemployment benefits (and hence the replacement rate) were increased by about 15 percent for workers earning below a certain threshold, whereas the replacement rate remained unchanged for workers above this threshold. The increase in the maximum duration of benefits depended on age and experience: for workers below age 40 and/or workers with little previous work experience, the duration remained unchanged; for workers with long previous work experience, the duration increased, from 30 to 39 weeks for the age group 40–49 and from 30 to 52 weeks for workers aged 50 and older. Rafael Lalive, Jan van Ours, and Josef Zweimüller used this natural experiment to derive lessons about the effects on unemployment duration of changes of UB systems along these different dimensions. An additional advantage of this reform is that the policy change took place under a stable macroeconomic environment, notably under a constant unemployment rate, which reduces the risk of endogenous policy bias.

The data analyzed concerned workers' unemployment histories from two years before the policy change to two years after the change. The authors estimated

(continued)

Box 11.4 (continued)

the conditional probability of leaving unemployment at different durations of unemployment and for workers entitled to different durations of UBs and replacement rates. The table in this box summarizes their main results by comparing the expected duration of unemployment (measured in weeks) in the first 104 weeks of the unemployment spell for groups involved by the various regulatory changes (the treatment group) and groups not involved (the control group).

	Treated	Control	Effect
No treatment	16.91	16.91	0.00
Change to one parameter			
PBD 30–39 weeks	17.53	17.08	0.45
PBD 30–52 weeks	20.62	18.35	2.27
RR increase	20.97	20.60	0.38
Change to two parameters			
PBD 30–39 weeks and RR increase	21.95	21.09	0.86
PBD 30–52 weeks and RR increase	29.43	23.70	5.72

They found that both the increase in replacement rates and the extension of benefit duration significantly increase the duration of unemployment, and that the effect is stronger when maximum duration is increased rather than the replacement rate or when replacement rates and benefit durations are simultaneously increased. In particular, extending the benefit duration by 9 weeks tends to increase expected unemployment duration by 0.45 weeks, or by 0.05 weeks per additional week of maximum duration (second row). Increasing benefit duration by 22 weeks generates about 2.27 additional weeks of unemployment (third row). In contrast, increasing the replacement rate by 6 percentage points tends to prolong unemployment duration by 0.38 weeks (fourth row). Finally, individuals eligible for a combined 9 week increase of the maximum duration of UBs and a 6 percentage point increase in replacement rates are unemployed for 0.86 weeks longer than in the counterfactual situation without this combined intervention (fifth row). Interestingly, this change in unemployment duration matches almost perfectly the prediction obtained from two separate changes; that is, $0.38 + 0.45 = 0.83$. A Stata data file with the Lalive, Van Ours, and Zweimüller dataset, a do file, and a logfile are available at the website <http://press.princeton.edu/titles/8771.html>.

Source: Lalive, Van Ours, and Zweimüller (2006).

11.3.4 Effects of Unemployment on the Generosity of UBs

How can the relatively strong effects of an increase of replacement rates on unemployment found by the macroeconomic literature be reconciled with the mild effects estimated from micro data? A possible explanation of this discrepancy has to do with *policy endogeneity* or *reverse causality*, whereby higher unemployment

induces governments to adopt more generous benefit systems rather than the other way around (Holmlund 1998). Estimates based on macroeconomic data in the presence of this reverse causality may also attribute to the generosity of UBs the effects of other factors (e.g., other labor market institutions) that increase the duration of unemployment.

Some studies have tried to isolate the effects of labor market conditions on UB generosity. Regional diversification in the maximum duration of UBs was found to be positively correlated with the duration of unemployment in U.S. states (Card and Levine 2000). Lalive, and Zweimueller (2002) showed that estimates that do not correct for this policy endogeneity may significantly overstate the negative effects of UBs on the duration of unemployment in Austria. In particular, they estimated that the effects of the increase in benefit duration from 30 to 209 weeks on unemployment duration would have been 40 percent larger without correcting for policy endogeneity.

The preceding discussion suggests that unless macroeconomic studies adopt proper methods to correct for the potential presence of policy endogeneity, it may be wiser to consider the estimates of the effects of UBs on employment and unemployment based on macroeconomic data only as upper bounds (Nickell 1997; Blanchard and Wolfers 2000).

11.4 Policy Issues

11.4.1 Why and When Should UBs Be Publicly Provided?

The theory and empirical evidence reviewed in this chapter suggest that UBs raise a worker's reservation wage and increase unemployment duration. UBs also need to be financed via payroll taxes, which, as discussed in chapter 4, are distortionary in that they reduce the size of the labor market. The obvious question is why unemployment insurance is publicly provided. The answer is that UBs provide insurance against an otherwise uninsurable unemployment risk. No private company around the world supplies insurance against the risk of becoming or remaining unemployed because moral hazard and adverse selection stand in the way of such potential contractual arrangements. *Moral hazard* arises because workers do not try as hard to avoid unemployment and find new jobs if they are covered against the negative consequences of the event by purchasing insurance at a given market price. *Adverse selection* arises in the presence of a heterogeneous pool of unemployed individuals, and one in which some workers have (unobservable) characteristics that make them more likely to experience a job loss. Under these conditions, workers who know that their unemployment risk is particularly high can make the scheme unprofitable for private insurance providers and force them to raise prices, making the insurance scheme unattractive to workers with average risk.

Public provision of UBs therefore remedies the inequitable or unfair labor market treatment of workers who, lacking insurance, become or remain unemployed despite their best efforts. Because UBs are *contingent* payments (occurring only in the case of job loss), they also bring about greater utility compared with individual savings. Public provision of UBs solves the adverse selection problem because participation in the insurance is compulsory for everybody. However, it does not solve the moral hazard problem because workers have no less incentive to decrease their job search intensity when covered by public rather than private insurance. Unemployed individuals will not be actively searching for other jobs if their utility is not much different when they work and when they do not work. The evidence reviewed in this chapter suggests that disincentives to job search associated with UB provision can indeed be substantial. Moreover, protection from unfair developments unavoidably decreases the labor market's speed of adjustment by increasing the duration of unemployment. More broadly, labor market institutions that protect workers against unfair market developments unavoidably reduce the intensity of competition because they trade off lower productive efficiency against *ex ante* distributional equity. This is desirable, from an *ex ante* point of view, in the presence of the informational asymmetries that prevent private insurers from doing the job, but *ex post* unavoidably decreases productive efficiency.

Such efficiency losses may not be easily affordable for developing countries, but are generally accepted by rich and relatively stable societies. While all OECD countries have a UB system in place, in developing nations unemployment benefits are either absent or poorly enforced, leaving a large fraction of the jobless population without income transfers. According to information collected by the U.S. Department of Labor (Social Security Programs Throughout the world), UBs are offered in 17 of the 27 Asian countries, in 8 of the 22 Central and Latin American economies, in just 4 of the 15 Arab countries (Algeria, Egypt, Iran, and Iraq), and in only 3 of the 32 African economies (Nigeria, South Africa, and Tunisia). Vodopivec (2004) also documents that in low-income countries income "support schemes for the unemployed consist almost exclusively of public works and severance pay programs." In middle-income countries unemployment benefits involve relatively high nominal replacement rates (e.g., 60 percent of the best earnings in the last year in Argentina), but are offered for a short period and are not fully indexed. Under rapid inflation, unemployment benefits tend to collapse to the benefit minima, generally established as a fraction of the minimum wage (which is itself low as a fraction of the average wage, by Western standards). Many middle-income countries do not have an unemployment insurance or unemployment assistance scheme. As documented by Vodopivec (2004), for instance, only one Latin American country out of five has a UB system, and only one out of three in Asia. Yet more and more countries are introducing UB systems (Boeri and Macis 2008).

11.4.2 What Is the Optimal Structure of UBs?

As argued earlier, UBs are multidimensional institutions. Hence proper design of UBs requires nontrivial fine-tuning of this institution along several dimensions, addressing different trade-offs simultaneously.

The first of these trade-offs was addressed in the previous section: it involves balancing the insurance, consumption-smoothing function of UBs with the efficiency losses associated with the potential reduction in job search efforts when insurance is provided. Because of these losses, optimal UBs should offer *incomplete* insurance against the risk of job loss: workers should also partly bear this risk. The appropriate degree of risk sharing between the state and the workers can be determined by solving an optimal insurance problem. The design of UBs described at the beginning of this chapter can be rationalized in terms of optimal insurance arguments. When search effort is unverifiable, and hence the principal (the state) must give to the agent (the unemployed individual) an incentive to make this effort, optimal unemployment insurance models (Hopenhayn and Nicolini 1997) prescribe earnings related UBs offering replacement rates that decrease over time. Another feature of the benefit system—the flat and open-ended benefit provided at longer unemployment durations—is consistent with prescriptions of economic theory. Because human capital depreciates during the unemployment spell, at longer unemployment durations, eliciting search effort may become too costly relative to offering support to nonemployed individuals (Pavoni and Violante 2007). Finally, the welfare-to-work schemes that combine passive income transfers with active labor market policies, which are characterized in chapter 12, all involve a sequence of actions and sanctions at different durations of unemployment.

Another trade-off related to the provision of UBs that has been more recently highlighted by the literature has to do with the *quality of jobs* and the effects of UBs on structural change. Acemoglu and Shimer (1999, 2000), as well as Marimon and Zilibotti (1999), who considered bilateral job search models with heterogeneous workers and jobs, showed that a positive level of UBs is required to maximize output because otherwise jobseekers, notably poorer jobseekers, would be discouraged from applying for high-productivity jobs that are more difficult to get, and firms would not create this type of jobs. In these models unemployment generated by UBs has an efficiency-enhancing function in that it allows economies to generate better matches and hence higher-productivity jobs. There is some evidence (Boeri and Macis 2008) that historically the introduction of UBs has indeed been associated with an increase in job reallocation and structural change. The experience of countries coming from central planning is also consistent with UBs increasing job reallocation and structural change (Boeri 2000). At the same time, when UBs are too generous, efficiency-enhancing job reallocation may be reduced because unemployment duration increases and stagnant unemployment pools are created.

Finally, unemployment benefits will have to be paid by someone. Thus the choice of the proper design of UBs involves also an *optimal taxation* problem. From the standpoint of workers, there is a trade-off between the higher contributions that they have to pay to the unemployment insurance scheme and the insurance that they obtain against unemployment risk. The trade-off is more favorable to paying higher contributions for workers with higher-than-average probability of job loss. Because contribution to UBs is mandatory, they redistribute in favor of workers more exposed to unemployment risk, even when neither contribution rates nor benefit rules are progressive. A higher elasticity of labor supply would also shift the burden of paying unemployment insurance onto employers, reducing the effects on net wages of the introduction of UBs. From the standpoint of society as a whole, the choice of the optimal UB tax will depend on the relative elasticity of labor supply and demand, as discussed in chapter 4.

11.5 Why Do Unemployment Benefits Exist?

An increasing number of countries in the world, even middle-income countries with serious fiscal constraints, are introducing unemployment benefits systems. Many European countries are providing more and more protection against income fluctuations due to job loss via UBs rather than EPL, along with the flexicurity arrangements discussed in chapter 13. This suggests that these schemes not only are popular among workers, who receive insurance against otherwise uninsurable risk, but also increase welfare from the standpoint of society as a whole. As hinted by recent developments of job search theory, properly designed UBs improve the allocation of human capital and thus, foster economic growth. This means that UBs should not be considered a sort of luxury good to be offered only by rich countries. At the same time, UBs should not be too generous in order not to discourage job search altogether and generate stagnant unemployment pools. The most relevant issues do not concern whether or not a country should have a UB system, but how the system should be designed along its several dimensions.

Under the pressure of workers' organizations, rich countries often end up offering too-generous UB systems, inducing long-duration unemployment. It is then politically difficult to reduce the disincentives to job search associated with UB provision. The inventory of reforms reviewed in chapter 1 suggests that governments find it hard to enforce a relatively short duration of benefits or reduce replacement rates in long spells of unemployment. These reforms are likely to be vetoed by those who have paid for UBs and have entitlements to more generous treatment under the existing rules. Indeed, UB reforms typically grandfather existing entitlements by exempting those who are receiving the benefit at the time of the reform and allowing for a gradual phasing-in of the new rules. In order to counter opposition to reforms, UBs are typically scaled down via a number of marginal adjustments of the benefit formula and a gradual tightening of entitlement rules.

This creates two-tier systems in which only a fraction of the workforce is under the new regime.

Overall, reforms that reduce the generosity of UBs are politically difficult. It is less costly for policymakers to introduce activation measures and financial incentives that increase job-finding rates of the unemployed without affecting the pure insurance function of UBs. Reforms are often carried out simply by changing the administration of UBs, for example, conditioning the receipt of benefits on more frequent visits to the labor office and changing the definition of a suitable job offer (the unemployed who refuse to take a suitable job may be sanctioned by an interruption of benefit payments). These rules are discussed in chapter 12. They are clearly less visible (because they do not have to go through a legislative process) and quite effective in reducing disincentives to job search. The empirical evidence reviewed in this chapter suggests that reducing the maximum duration of benefits has the potential to reduce the duration of unemployment significantly, but that this is not necessarily the case with reductions of replacement rates. Reforms also combine benefit cuts with employment-conditional incentives or wage subsidies that increase job-finding rates. This is likely to win support from employees and reduce the opposition of the unemployed. These activation schemes are also discussed in the next chapter.

11.6 Suggestions for Further Reading

It is time to write a new survey of the rich empirical literature on the effects of UBs on unemployment duration. Although existing surveys are a bit dated, they can still be very instructive. We recommend in particular the surveys by Tony Atkinson and John Micklewright for the *Journal of Economic Literature* (1991) and James Devine and Nicholas Kiefer (1991), the survey of results from duration analysis offered by Nicholas Kiefer (1988), and the more recent surveys by Bruce Meyer (1995) and Alan Krueger and Bruce Meyer (2002), which, however, cover mainly the U.S. experience.

11.7 Review Questions

1. Why do replacement rates offer an incomplete measure of the generosity of unemployment benefits?
2. Do unemployment benefits redistribute in favor of low-skilled workers?
3. What effects should we expect UBs to have on first-time jobseekers when benefits are conditional on previous work experience?
4. How does the introduction of a UB system affect labor force participation?
5. What is the hazard rate?

6. What type of relation do we expect to observe between generosity of UBs and structural change?
7. Why is unemployment insurance not provided by private insurance companies?
8. Why is there a socially optimal replacement rate?
9. Should UBs be experience rated?

11.8 Technical Annex: UBs in a Search-Matching Model

11.8.1 Effects of UBs in a Two-Sided Job Search Model

Unemployment benefits (UBs) aim at protecting workers against uninsurable labor market risk. The generosity of UBs is determined by all the different features of this institution, for example, eligibility conditions, duration of benefits, and amount of transfers. Nevertheless, in this simplified two-sided job search model we treat unemployment benefits as a one-dimensional institution where only the replacement rate b matters and benefits last indefinitely.

As discussed in the main text, changes in the generosity of unemployment benefits are bound to affect search intensity, wage bargaining, and labor force participation. In this annex we formally characterize these effects. Because there are also important feedback, second-round effects, we also present some numerical simulations of the model that are useful in gauging the overall effects of an increase in UBs. The model is an extension of the model presented in technical annex 10.8, allowing for an endogenous search intensity, taxes (to pay for unemployment benefits), and, of course, unemployment benefits themselves.

11.8.2 Optimal Job Search Intensity

All workers are assumed to be homogeneous so that they receive the same wage w net of taxes. Unemployed workers search with intensity s . A worker can adopt search intensity $s \in [0, 1]$ in looking for a job. The disutility of searching at intensity s equals $\gamma(s)$, with $\gamma(0) = 0$, $\gamma'(0) = 0$, $\gamma'(s) > 0$, $\gamma''(s) > 0$, and $\lim_{s \rightarrow 1} \gamma'(s) = +\infty$. The following search cost function is used:

$$\gamma(s) = \frac{1 - (1 - s)^{1-\kappa}}{1 - \kappa} - s, \quad (11.7)$$

with $0 < \kappa < 1$.

Unemployed workers receive unemployment benefits $b = b^w$, where b^w is the replacement rate and $0 < b^w < 1$. The flow value of being unemployed equals the difference between the cost of search and the unemployment benefits and the

expected benefits from finding a job:

$$\rho V_u = \max_{0 \leq s \leq 1} [b - \gamma(s) + \mu s (V_e - V_u)], \quad (11.8)$$

where μs is the job-finding rate and the optimal search intensity s^* is found through first derivatives:

$$\gamma'(s^*) = \mu(V_e - V_u). \quad (11.9)$$

The flow value of having a job is

$$\rho V_e = w + \delta(V_u - V_e), \quad (11.10)$$

where δ is the exogenous job separation rate.

11.8.3 Vacancies and the Matching Function

Vacancies and unemployed workers meet and match. This process is described by a matching function that, for a given flow of matches, is equivalent to the so-called Beveridge curve:

$$m(su, v) = A(su)^{1-\eta} v^\eta. \quad (11.11)$$

The matching process is specified as a Cobb-Douglas function with constant returns to scale and A and η as relevant parameters. The matching function can be rewritten in terms of μ , the search parameter for unemployed workers:

$$\mu = \frac{m(su, v)}{su} = A\theta^\eta, \quad (11.12)$$

where θ is labor market tightness—the number of vacancies divided by the effective number of unemployed:

$$\theta = \frac{v}{su}. \quad (11.13)$$

Given that production is sold at (numeraire) price 1, the flow value of a job is

$$\rho J_e = y - (1 + t)w + \delta(J_v - J_e), \quad (11.14)$$

where t is the tax rate. The flow value of a vacancy is

$$\rho J_v = -c_v + \frac{\mu}{\theta}(J_e - J_v), \quad (11.15)$$

where $\frac{\mu}{\theta}$ is the rate at which vacancies are filled and c_v is the cost of an unfilled vacancy.

11.8.4 Wage Bargaining

To determine the wage, the Nash-bargaining solution (see technical annex 3.8) is used. The bargaining power of the worker is denoted by $\beta \in [0, 1]$ and the bargaining power of the employer by $(1 - \beta)$. Then the net wage w solves

$$\max_w (V_e - V_u)^\beta (J_e - J_v)^{1-\beta}, \quad (11.16)$$

or

$$\max_w \left(\frac{w}{\rho + \delta} + \frac{\delta}{\rho + \delta} V_u - V_u \right)^\beta \left(\frac{y - (1+t)w}{\rho + \delta} + \frac{\delta}{\rho + \delta} J_v - J_v \right)^{(1-\beta)}. \quad (11.17)$$

The first-order condition for w can now be written as

$$(1 - \beta)(1+t)(w - \rho V_u) = (1 - \beta)(y - (1+t)w), \quad (11.18)$$

or, equivalently,

$$w = (1 - \beta)\rho V_u + \beta \frac{y}{1+t}. \quad (11.19)$$

The higher the worker's bargaining power (the higher the β), the higher the share of the after-tax surplus $\frac{y}{1+t}$. In case workers have full bargaining power ($\beta = 1$), the wage is equal to the total after-tax surplus. In case employers have full bargaining power ($\beta = 0$), the wage is equal to the flow value of unemployment ($w = \rho V_u$). In the first case the worker gets the entire surplus. In the second case the employer extracts all the surplus, conditional on the worker having some incentive to work.

11.8.5 Labor Market Participation

In a steady state labor market there are constant stocks of unemployed workers and vacancies. For unemployment this implies that the inflow into unemployment equals the outflow from unemployment:

$$\mu su = \delta e, \quad (11.20)$$

where e is employment. Furthermore, individuals are assumed to have heterogeneous preferences ω with respect to the value of leisure that are unobservable by the employer and are distributed according to $G(\omega)$ in the range $[\underline{\omega}, \bar{\omega}]$, such that $G(\bar{\omega}) = 1$. Then, if we assume that the total population of the relevant age is 1, the participation rate is

$$e + u = G(\rho V_u). \quad (11.21)$$

Note that because ω is unobservable, preferences for leisure do not affect wage bargaining. So, whereas individuals are heterogeneous with respect to their preferences for leisure, they are homogeneous once they enter the labor market.

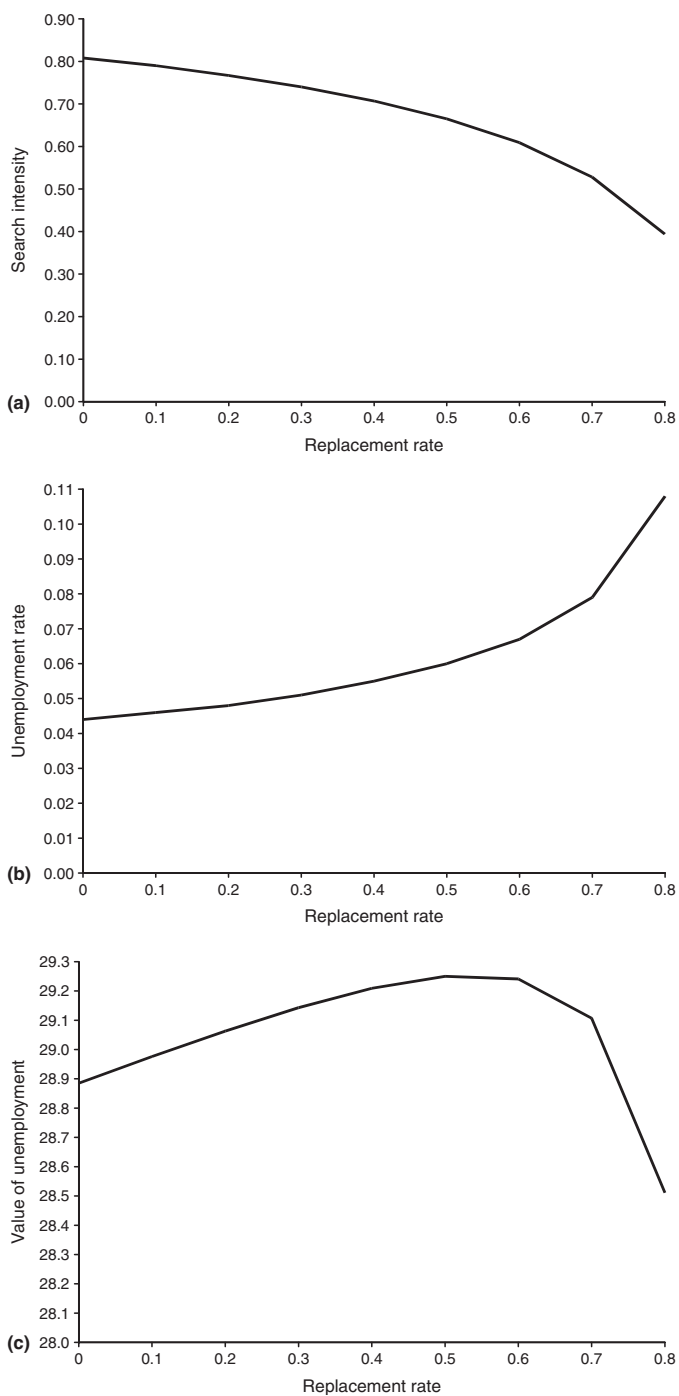


Figure 11.5 Effects of Replacement Rate on (a) Search Intensity, (b) Unemployment Rate, and (c) Value of Unemployment

11.8.6 Balanced Budget

We assume that the unemployment benefit level replacement rate b^w is defined as a fraction of the economy-wide average wage w . Then an individual worker's choice of wage w does not affect the unemployment benefit level that the worker will receive when that worker becomes unemployed. That is, $\frac{\partial V_u}{\partial w} = 0$ in the Nash-bargaining function. The presence of unemployment benefits implies that taxes have to be introduced. Taxes are to pay only for unemployment benefits. The marginal wage tax t is adjusted endogenously in the simulations to satisfy the government budget constraint

$$etw = ub^w. \quad (11.22)$$

11.8.7 Numerical Simulations

The model for simulation has 15 endogenous variables, $\gamma, s, V_u, V_e, w, J_e, J_v, v, u, e, m, \theta, \mu, t$, and $F(\rho V_u)$, and 9 parameters for which the following baseline values are used: replacement rate $b = 0$, discount rate $\rho = 0.025$, search cost function $\kappa = 0.5$, matching function $\eta = 0.5$, $A = 1.0$, wage negotiations $\beta = 0.5$, productivity $y = 1$, vacancy costs $c_v = 2$, and job destruction rate $\delta = 0.04$. Furthermore, for simplicity it is assumed that $G(\rho V_u) = 1$. Finally, since there is free entry in posting vacancies, $J_v = 0$.

Figure 11.5 shows how unemployment benefits affect search intensity, unemployment rates, and the value of unemployment. Search intensity decreases at an increasing rate if the replacement rate goes up (top panel), causing the unemployment rate to increase rapidly (middle panel). The relationship between replacement rate and the value of unemployment is nonlinear (bottom panel of figure 11.5). The replacement rate has a positive direct effect on the value of unemployment because it increases income while a person is unemployed. However, the replacement rate has a negative indirect effect on the value of unemployment: through the decline in search intensity and the decrease in labor market tightness, the transition rate from unemployment to employment decreases. At lower replacement rates the positive effect of the replacement rate dominates, but at higher rates the negative effect dominates. As shown for the parameter values of the current model, the maximum value of unemployment is at a replacement rate of about 0.5.

This page intentionally left blank

Active Labor Market Policies

Active labor market policies (ALMPs) have a long-standing tradition in many countries. At the beginning of the twentieth century employment offices were built up. In the depression of the interwar years government programs were established to put the unemployed to work. Later, labor market retraining was organized to stimulate occupational and regional mobility to facilitate structural adjustments.

Currently ALMPs aim at improving the functioning of the labor market by enhancing labor market mobility and adjustment, facilitating the redeployment of workers. ALMPs intend to overcome market failures arising from generous unemployment benefit and welfare benefit schemes. There are four basic functions of ALMPs: to raise output and welfare by putting the unemployed to work or having them invest in human capital, maintain the size of the effective labor force by keeping up competition for available jobs, help reallocate labor between different submarkets, and alleviate the moral hazard problem of unemployment insurance (Calmfors 1995).

There are four main types of ALMPs: (1) training, (2) subsidized employment, (3) public employment services (PESs), and (4) activation. *Labor market training* concerns training for unemployed adults, those at risk of losing their jobs, and employed adults. *Subsidized employment* consists of targeted measures to promote or provide employment for the unemployed and other priority groups. It also concerns wage subsidies paid to private-sector firms in order to encourage the recruitment of targeted workers or continued employment of those whose jobs are at risk, and support of unemployed persons starting enterprises, as well as direct job creation in public or nonprofit organizations for the benefit of the unemployed. PESs concern placement, counseling and vocational guidance, job search courses, and administration of unemployment benefits. ALMPs also include *activation*, measures that provide incentives for the unemployed to increase job finding either directly through benefit sanctions or through mandatory participation in training or subsidized employment. Key examples of activation programs are requirements that unemployed individuals attend intensive interviews with employment counselors, apply for job vacancies as directed by employment counselors, independently search for job vacancies and apply for jobs, accept offers of suitable

work, participate in the formulation of an individual action plan, and attend training or job creation programs. If unemployed workers are unwilling to participate in the activation programs, they may lose their benefits entitlement permanently or temporarily (in case of benefit sanctions). Some activation programs are *workfare* in the sense that they do not deliver further services except for keeping the unemployed busy. The main motivation for workfare schemes is that they enable a distinction between involuntarily and voluntarily unemployed persons. Most OECD countries' activation strategies in principle aim not to use workfare: all programs are intended to have employment service functions (OECD 2005b).

ALMPs may eliminate mismatch in the labor market, promote more active search behavior on the part of jobseekers, and have a screening function because they substitute for regular work experience in reducing employer uncertainty about the employability of job applicants. Placements in labor market programs may provide a work test as an alternative to eligibility for unemployment benefits, since some of those who are not genuinely interested in work will prefer to lose registration rather than participating in a program. An adverse side effect of ALMPs is that workers are locked in to training and job creation programs: because of their participation, they reduce their search intensity.

ALMPs are part of a comprehensive strategy to help transitions from unemployment or inactivity to work and smooth transitions from job to job. The behavior of unemployed individuals is also influenced by payroll tax schemes and in-work benefits (chapter 4) and unemployment benefits (chapter 11).

12.1 Measures and Cross-Country Comparisons

ALMPs are often but not always expensive. Depending on the amount of money, there may be few workers in expensive programs or many workers in cheap programs. Therefore, to indicate the importance of ALMPs at a country level, two measures are used. First, the number of workers participating in ALMPs as a share of the labor force; second, the amount of money spent on ALMPs as a percentage of GDP.

There are large cross-country differences in the share of the labor force participating in ALMPs, in 2005 ranging from a low of 1.2 percent in the Czech Republic to a high of 7 percent in Belgium. As shown in table 12.1, expenditures on ALMPs measured as a percentage of GDP differ greatly between countries. Whereas Mexico only spends 0.01 percent of GDP on labor market training, Denmark spends 0.51 percent. While Korea spends 0.03 percent of GDP on PES and countries like Italy, Luxembourg, and the United States devote little more to the public employment service, the United Kingdom spends 0.38 percent of GDP on the PES. With respect to expenditures on subsidized jobs, the differences between countries are also quite large. Italy, Korea, Mexico, New Zealand, the United

Table 12.1 Active Labor Market Policies: Participation (Percentage of Labor Force), Public Expenditures (Percentage of GDP), and Sanction Rates (Percentage of Unemployment)

	Active labor market policies						Sanction rates
	Labor force involved (%)	Public expenditures (percentage of GDP)					
		Total	Training	PES	Job creation	Other	
Australia	1.8	0.45	0.04	0.26	0.08	0.07	3.3
Austria	—	0.62	0.33	0.17	0.04	0.08	—
Belgium	7.0	1.08	0.20	0.23	0.39	0.26	0.8
Canada	2.5	0.32	0.08	0.16	0.02	0.06	6.1
Czech Republic	1.2	0.25	0.01	0.12	0.03	0.11	—
Denmark	5.2	1.74	0.51	0.32	—	—	2.1
Finland	3.7	0.89	0.37	0.20	0.07	0.25	10.2
France	3.6	0.90	0.29	0.25	0.18	0.18	—
Germany	4.7	0.97	0.25	0.35	0.10	0.27	1.1
Greece	—	—	0.03	—	—	—	—
Hungary	—	0.29	0.04	0.09	0.06	0.10	—
Ireland	3.6	0.63	0.24	0.12	0.21	0.06	—
Italy	—	0.54	0.20	0.08	0.01	0.25	—
Japan	—	0.25	0.04	0.19	—	—	0.02
Korea	—	0.13	0.04	0.03	0.01	0.05	—
Luxembourg	—	0.52	0.13	0.06	0.13	0.20	—
Mexico	—	0.02	0.01	—	0.01	—	—
Netherlands	4.2	1.33	0.13	0.49	0.18	0.53	36.0
New Zealand	1.9	0.39	0.17	0.12	0.01	0.09	0.4
Norway	2.7	0.75	0.37	0.12	0.07	0.19	7.3
Poland	—	0.43	0.10	0.07	0.03	0.23	—
Portugal	—	0.69	0.29	0.17	0.03	0.20	—
Slovak Republic	5.5	—	0.02	—	0.06	—	—
Spain	—	0.78	0.17	0.13	0.09	0.39	—
Sweden	4.4	1.32	0.34	0.23	—	—	0.6
Switzerland	2.5	0.76	0.29	0.14	—	—	38.5
United Kingdom	—	0.49	0.09	0.38	0.01	0.01	5.5
United States	—	0.13	0.05	0.03	0.01	0.04	35.4

Sources: OECD (2006a); Boone and Van Ours (2004); Grubb (2000).

Notes: Active labor market expenditures: 2005; sanctions and benefit refusals for behavior during benefit periods as a percentage of the average stock of benefit claimants, 1997–1998. The numbers refer to sanctions for labor market behavior conditions (not to administrative infractions).

States and the United Kingdom spend 0.01 percent of GDP on subsidized jobs, while Belgium spends 0.4 percent of GDP on this ALMP. The evolution of ALMP spending is also markedly different across countries. In Denmark, for example, public expenditures on labor market training went up from 0.4 percent in the second half of the 1980s to 0.99 percent in the second half of the 1990s. At the same time, in Ireland these expenditures went down from 0.5 percent to 0.2 percent of

GDP. In Sweden the expenditures first went up from 0.5 percent in the second half of the 1980s to 0.8 percent in the early 1990s and then down to 0.6 percent in the second half of the 1990s. At the lower end of the distribution, Japan kept expenditures on labor market training constant at a level of 0.03 percent of GDP. Similar differences occur with respect to public expenditures on subsidized jobs. In Belgium these expenditures were always at a high 0.4 percent of GDP, while in Ireland and Sweden there was a major increase from about 0.3 percent in the second half of the 1980s to 0.9 percent in the second half of the 1990s.

Finally, table 12.1 provides an overview of sanction rates (expressed as a percentage of the stock of unemployment benefit claimants) for some of the OECD countries. The sanctions refer to behavior during benefit periods.¹ The sanction rates range from very low in Belgium, Japan Sweden and New Zealand to quite large in the Netherlands and Switzerland. The Swedish system is sometimes considered one where there is pressure on the unemployed, including possible denial of benefits, both to look for work and to accept suitable job offers. Nevertheless, Björklund and Holmlund (1991) report that yearly benefit denials amount to no more than approximately 1 to 2 percent of all those who receive unemployment compensation during a year, and this rate has recently been falling. Thus, whereas the Swedish system is known for its active labor market policy, in terms of benefit sanctions it is the least strict of all the OECD countries represented in table 12.1. In principle, there are two reasons why observed sanction rates in a country are low. First, it may be that sanctions are seldom imposed because the system is lax and not credible. Second, a low sanction rate can be due to the (equilibrium) reaction by workers to comply with the search guidelines. In the case of Sweden, there are indications that the first reason is the cause. Since the sanction is a 100 percent reduction in the unemployment benefit, monitors are very reluctant to impose this sanction (see Björklund and Holmlund 1991 for details).

12.2 Theory

Figure 12.1 illustrates how ALMPs might affect the functioning of a labor market using the well-known Beveridge curve (box 12.1), the empirical relationship between unemployment rates and vacancy rates (unfilled vacancies as a percentage of the labor force). Curve BC_1 in figure 12.1 slopes downward because in a boom there are many vacancies and few unemployed, while in a slump there are many unemployed and few vacancies. As an economy experiences cyclical fluctuations, the unemployment-vacancy combination moves up and down the curve BC_1 . At a particular moment it may be that the labor market has an unemployment rate of u_1 and a vacancy rate of v_1 . The exact location of the Beveridge curve depends on

¹ Other sanctions may concern a lack of effort to prevent job loss (voluntary unemployment). In terms of the flow of initial benefit claims, these sanctions range from 3.4 percent in Finland to 13.5 percent in the United States See Grubb (2000) for details on the system of benefit sanctions in various countries.

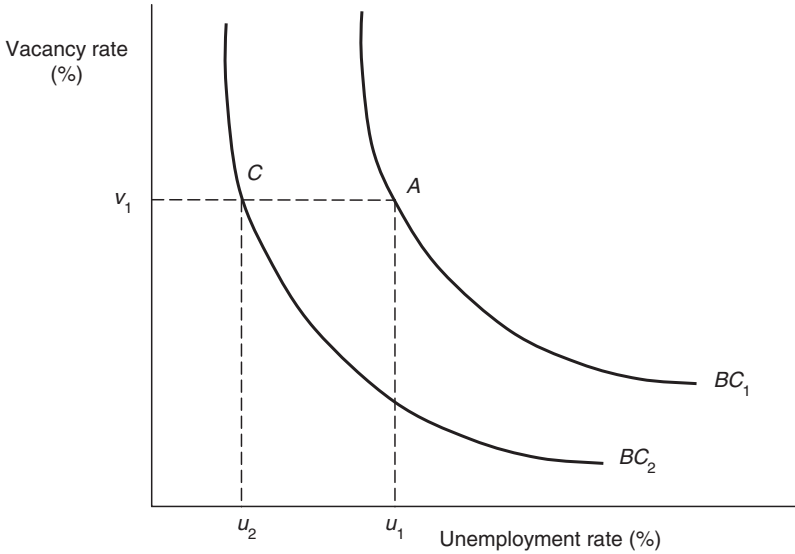


Figure 12.1 The Beveridge Curve

how efficient the labor market is. ALMPs may improve the efficiency of the labor market, causing an inward shift of the Beveridge curve, for example from BC_1 to BC_2 . If that is the case, conditional on the vacancy rate v_1 , the unemployment rate will go down from u_1 to u_2 .

Box 12.1 The Beveridge Curve and the Matching Function

In many labor markets there appears to be an inverse relationship between unemployment rates and vacancy rates. The graphical representation of this relationship is named the Beveridge curve after William Beveridge (1879-1963), a British economist who was the first to note this empirical regularity (in 1942). Later, theoretical foundations for the Beveridge curve were provided in which the labor market was described by a matching process (Pissarides 1979; Blanchard and Diamond 1994). Through this matching process vacancies are filled and unemployed workers find jobs. The process may be described by a matching function:

$$m = A(U)^{1-\alpha}V^\alpha,$$

where m is the number of matches per time period, $U(V)$ is the number of unemployed workers (vacancies) at the beginning of that time period, A represents the efficiency of the matching process, and α is usually assumed to be equal to 0.5. In a dynamic labor market individuals change position frequently. Not only do unemployed workers find jobs, but also many employed workers lose their jobs

(continued)

Box 12.1 (continued)

because firms reduce their workforce temporally or permanently. The reallocation of labor across different firms causes a regular flow of workers into and out of unemployment. Often the flow of workers from employment to unemployment is assumed to be a constant fraction of total employment:

$$F_{in}^u = \delta^s L,$$

where F_{in}^u is the inflow into unemployment, δ^s is the job separation rate, and E is the number of employed workers. In a steady state labor market there are constant stocks of unemployed workers and vacancies; the inflow into unemployment F_{in}^u equals the outflow from unemployment m , and therefore

$$\delta^s L = AU^{1-\alpha} V^\alpha,$$

or

$$\frac{\delta^s}{A} = \left(\frac{U}{L}\right)^{1-\alpha} \left(\frac{V}{L}\right)^\alpha \approx u^{1-\alpha} v^\alpha,$$

where u is the unemployment rate and v is the vacancy rate. If δ^s and A are constant, there is a stable hyperbolic relationship between the unemployment rate and the vacancy rate, the Beveridge curve as depicted in figure 12.1. If the matching process becomes more efficient, for example, through ALMPs, the parameter A increases and the Beveridge curve shifts inward.

Boone and Van Ours (2004) present a theoretical search-matching model in which they distinguish three types of ALMPs: training of unemployed workers, subsidized employment, and public employment services. They model training as a subsidy to training costs of unemployed workers. The idea is that placement workers help unemployed workers find the most suitable courses for them so that they do not waste time and effort enrolling in less effective courses. Furthermore, the government sometimes creates courses that are directly relevant to targeted groups of the unemployed. This is less costly for the unemployed than doing bits and pieces from different courses, one of which is targeted at them. The effect of employment services is modeled as a subsidy to search costs of workers. Here the placement workers help filter all vacancy information so that only the vacancies most relevant for an unemployed worker come up. This reduces the search cost for the unemployed. Finally, subsidized employment is modeled as a subsidy to the value of the match of low-productivity jobs. In the model of Boone and Van Ours there are two channels through which ALMPs can potentially reduce unemployment. First, the job-finding rate may be increased. Second, through training the unemployed can get better jobs (with higher wages and lower job destruction rates). If ALMPs cause more unemployed workers to end up in high-skilled jobs, this reduces unemployment by decreasing the

flow from employment to unemployment. It turns out that the effects of ALMPs on the job-finding rate are theoretically ambiguous. However, the mechanism via the quality of the job and the flow from employment to unemployment discriminates between the different ALMPs. Boone and Van Ours (2004) show that training may do very well in reducing unemployment, while subsidized jobs and public employment services may not be as effective in reducing unemployment. They also show that there may be an interaction effect of unemployment benefits and training: training is more effective if unemployment benefits are more generous.²

Boone and Van Ours (2006) present a theoretical framework to analyze the impact of benefit sanctions, an activation measure to increase job-finding rates of recipients of unemployment benefits. They present a search-matching model in which benefit sanctions affect the intensity with which unemployed workers search for a job. Benefit sanctions may affect unemployment duration through two channels: the first operates *ex post*, and the second *ex ante*. Benefit sanctions will increase the search intensity of the sanctioned because of the reduction in the value of being unemployed. This effect is the *ex post effect*: an actual benefit reduction stimulates a worker's search efforts. Furthermore, the nonsanctioned may also increase search intensity because of stricter enforcement of job search requirements. This second effect is the *ex ante effect*: the risk of getting a benefit sanction influences the search behavior also of unemployed workers who have not been sanctioned. To illustrate how penalties and monitoring intensity affect the labor market, we performed some simulations of which the results are shown in the graphs in figure 12.2 (see technical annex 12.8 for details).

The upper graph in figure 12.2 shows how the size of the penalty affects unemployment rates. Increasing the penalty has an *ex ante* effect (workers are more eager to avoid being penalized and hence search harder) and an *ex post* effect (once penalized, they gain more from accepting a job and therefore search harder). The size of the penalty effect becomes smaller if the penalty increases. However, even if it is equal to 100 %, the unemployment rate will still be above the level it would have been in the absence of benefits. After all, even with a 100 % penalty, the unemployed receive benefits before they get a sanction imposed. The lower graph of figure 12.2 shows how the intensity of monitoring affects unemployment rates. The effect of an increase in monitoring intensity is similar to the effect of an increase in the penalty. If the monitoring rate goes up, search intensity will increase to avoid being penalized; after a sanction is imposed, search intensity may be lower than before because through the increase in monitoring intensity the value of being unemployed is reduced, and therefore the value of having a job is also lower.

² See Bassanini and Duval (2006) for similar results.

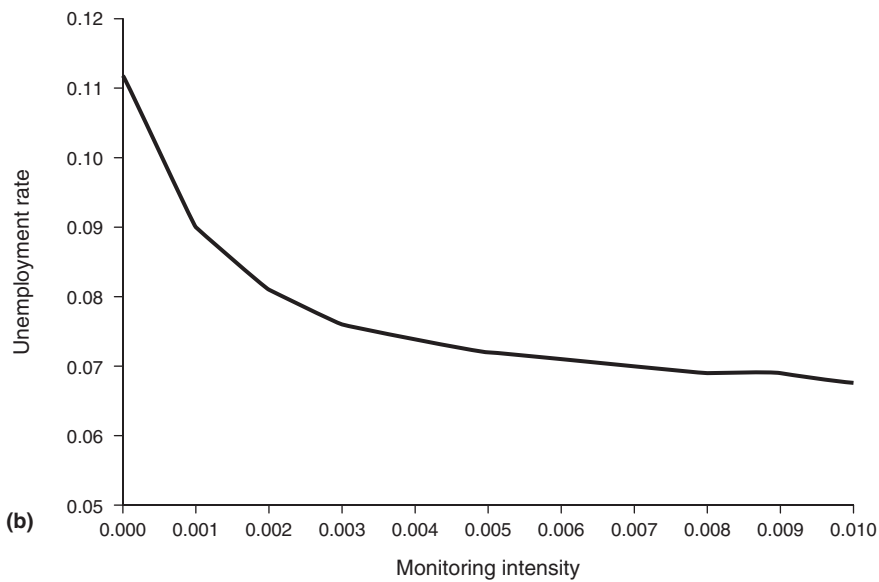
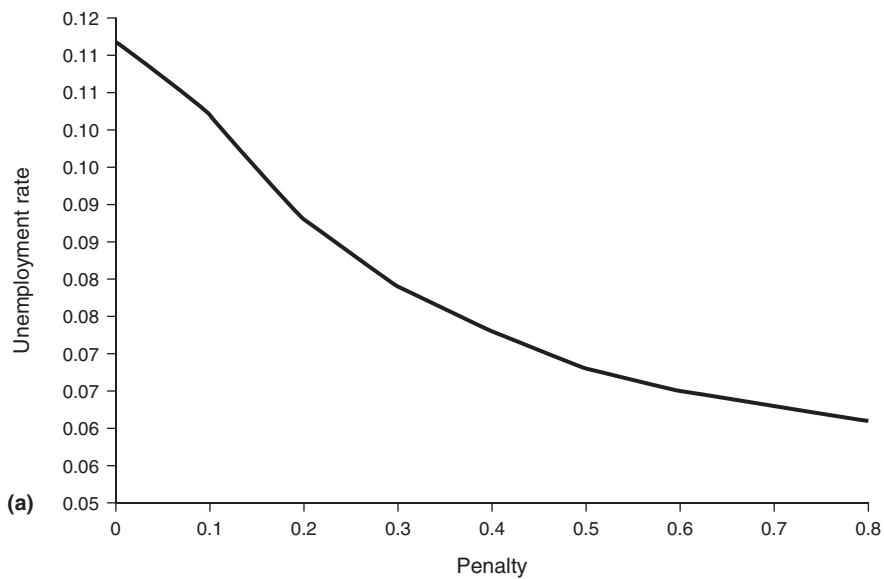


Figure 12.2 Benefit Sanctions and Their Effect on the Unemployment Rate: Effects of Penalty (a) and Monitoring Intensity (b)

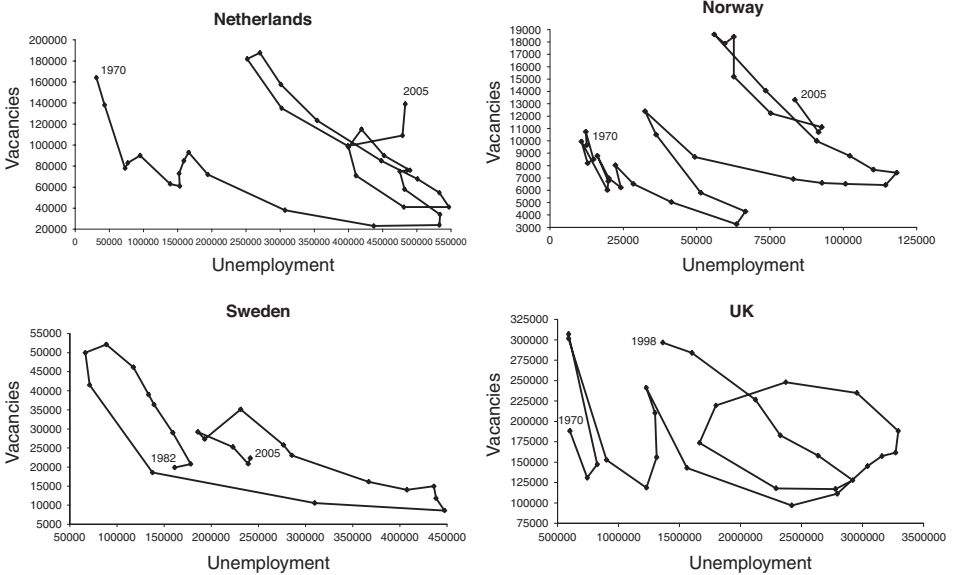


Figure 12.3 The Beveridge Curve in Four Countries

12.3 Empirical Evidence

Figure 12.3 provides empirical examples of the Beveridge curves for the Netherlands, Norway, Sweden, and the United Kingdom.³ In each of these countries the relationship between unemployment and vacancies has a hyperbolic curvature. There are also fluctuations and shifts outward of the curve, indicating a deterioration of the functioning of the labor markets.

However, it is difficult to draw conclusions on the effectiveness of ALMPs on the basis of shifts in the Beveridge curves. This is a general problem since the effectiveness of ALMPs is not obvious. Not only are there direct effects of ALMPs, but also there are indirect effects that can be quite important (Calmfors 1994). First, there are *displacement effects* because jobs created by one program can replace other jobs. Second, there are *deadweight effects* because labor market programs subsidize hiring that would have occurred anyway in the absence of the program. Third, there are *substitution effects* because jobs created for a certain category of workers replace jobs for other categories as wage relativities change. Finally, there are the effects of taxes (*fiscal substitution effects*) required to finance the programs on the behavior of everyone in society.

³ Many OECD countries do not have long series of vacancy data, which makes it difficult to generate Beveridge curves for more than a few countries.

12.3.1 Experimental Studies

Some studies exploit an experimental setup to address the effectiveness of ALMPs. Gorter and Kalb (1996) analyze the effects of intensive counseling and monitoring. By way of experiment, in seven employment offices in the Netherlands some unemployed workers were exposed to more intensive counseling and monitoring than others who got the usual treatment. The unemployed themselves were unaware of being part of the experiment. The allocation to more intensive counseling and monitoring was done randomly and at the start of the unemployment spell; the only way out of the experiment was through finding a job. Both the unemployed in the treatment group and in the control group had monthly meetings with staff of the employment office. The control group had traditional meetings in which progress in finding a job was discussed and occasionally the information provided by the unemployed worker was checked. The treatment group had longer meetings because the applications were more thoroughly discussed and there was more advice on potentially suitable vacant jobs and direction of search. Furthermore, the employment office staff spent more time checking information given by the unemployed. People who did not make sufficient applications and/or listed false applications to disguise their lack of search effort had a larger probability of being detected and penalized accordingly. Gorter and Kalb find that the treatment group had a higher application rate and a higher job-finding rate with a somewhat smaller matching probability—for the unemployed who previously had a permanent job, the application rate increased by 20 percent, while the job-finding rate increased by 15 percent. Thus more intense counseling and monitoring stimulated unemployed workers to make more applications for a job. Because of this, the job-finding rate increased. This shows that the additional applications were not fake applications.

Dolton and O'Neill (1996) present an analysis of the British Restart program, which consisted of a series of compulsory interviews every six months for unemployed workers starting after they have been registered as unemployed for six months. During this mandatory interview the counselor assessed the recent unemployment history of the worker, offered advice on search behavior and training courses, and sometimes initiated direct contact with employers. Unemployed workers were randomly assigned to the treatment group or the control group, whose members were eligible but did not have to attend the first interview. Those who were assigned to the treatment group were faced with the possibility of having their benefits reduced if they did not attend the Restart interview or if they were considered not to be making sufficient effort to find a job. Dolton and O'Neill find that the Restart interviews significantly decreased unemployment durations. Some individuals left the unemployment registers without having found a job; this was common among women and other groups "who are most likely not to be genuinely available for work." This type of outflow from unemployment was particularly high around the timing of the first Restart interview, which indicates a threat effect. However, measured over a period of 18 months, exits to a job were also

significantly different for the treatment group and the control group. In a follow-up article Dolton and O'Neill (2002) investigate the long-run effects of the Restart program and find that the Restart interviews reduced the male unemployment rate five years later by 6 percentage points compared with a control group for whom participation in the first interview took place six months later.

Klepinger, Johnson, and Joesch (2002), who study the effects of alternative work search requirements in Maryland, find that imposing additional search requirements speeds up the job-finding process (see box 12.2 for details). Black et al. (2003) present an analysis of an experiment from Kentucky. Unemployed workers were ranked in 20 categories according to a profiling score based on the expected unemployment duration. Local budgets available for reemployment services were allocated to the unemployed according to profiling scores, starting with the highest score, that is, the longest expected duration. In case there was insufficient funding for all the unemployed and the marginal group could not be covered completely either, there was a random allocation of reemployment services to the unemployed in the marginal group. By comparing the treated with the nontreated, the authors find that the reemployment services stimulated workers to leave unemployment more quickly; those who left unemployment quickly did not earn lower wages, which suggests that there is “no long-term harm from the treatment provided by the program.” The main action caused by the mandatory reemployment services was at the beginning of the unemployment spell. The unemployed who were notified of their obligations to attend the reemployment service program left unemployment quickly to avoid having to enter the program. In other words, the threat effect was driving the results. Apparently, many unemployed workers consider the reemployment service programs sanctions that they prefer to avoid.

Box 12.2 ALMPs in the United States

Daniel Klepinger, Terry Johnson, and Jutta Joesch present the results of an experimental evaluation of alternative work search requirements imposed on unemployment benefit recipients in Maryland. Assignment to control and treatment groups was random, based on Social Security numbers. The control group of unemployed workers had the standard obligation to contact two employers per week and report those contacts in order to remain eligible for unemployment benefits payments. The Maryland experiment distinguished four treatment groups who were informed about their duties within one week after registering for a benefit claim. The first group had to make four employer contacts per week, the second group was informed that they had to search actively without specifying the number of contacts they had to make, and the third group had to attend a four-day job search workshop lasting 16 hours early in the unemployment spell. The fourth group was informed that their claimed employer contacts would be verified. As discussed in the main text, there are two effects of increased work

(continued)

Box 12.2 (continued)

search requirements. First, the treatment effect: the unemployed may make more job contacts, which increases the job-finding rate. Second, the threat effect: the additional requirements raise the nonmonetary costs of remaining unemployed, which leads to more intense job search or a reduction in reservation wages (or both). Comparing the four groups allows a distinction between the two effects. The results for various outcome measures in the first year after the start of the unemployment spell are as follows:

Outcome Measure	Control group	Treatment group effects			
		Additional contacts	No reporting of contacts	Workshop	Verify contacts
Total UI benefits paid (\$)	2,085	−116*	34	−75*	−113*
Weeks of benefits	11.9	−0.7*	0.4*	−0.6*	−0.9*
Exhausted benefits (%)	28.3	−2.5*	1.5*	−1.1	−2.8*
Percentage worked	80.0	1.1	0.8	−0.8	1.3
Earnings (\$)	8,407	54	347*	−163	124

Note: * means significantly different from the control group at the 5 percent level.

The nonmonetary costs of imposing additional search requirements are important for the duration of benefit claims. Increasing the required weekly number of employer contacts from two to four and indicating that employer contacts would be verified reduced the duration of unemployment benefit spells by almost a week, which is a substantial effect since the average unemployment benefit duration was about 12 weeks. Eliminating a specific number of required contacts increased unemployment benefit duration. Finally, the obligation to attend a job search workshop also reduced unemployment benefit duration. This was at least partly because many unemployed workers left unemployment shortly before their search workshop was planned. Effects on the quality of postunemployment jobs in terms of employment and earnings were small or absent.

Source: Klepinger, Johnson, and Joesch (2002).

Van den Berg and Van der Klaauw (2006) analyze the outcome of a small-scale experiment on counseling and monitoring in the Netherlands. They distinguish between two types of job search, formal and informal job search. Formal job search is search through personnel advertisements and public employment offices. Informal job search refers to direct contacts with employers and search through friends, relatives, or employed workers. They argue that counseling and monitoring only affects formal job search. Monitoring leads to a substitution of effort from informal to formal search, which reduces the effectiveness of monitoring. At the time of the experiment in the Netherlands, at the start of the unemployment spell unemployed workers were classified into one of four types on the basis of objective characteristics and subjective evaluation. Types 2, 3, and 4 unemployed are offered assistance to find a job, while type 1 unemployed are expected to

have sufficient skills to find a job without assistance. The experiment was limited to type 1 unemployed, who were randomly assigned to a treatment group and a control group. The control group had to report on search activities every week, while in addition to this requirement the treatment group had regular meetings with counselors during which initially quality of application letters and résumés were examined and a plan was made, while during follow-up meetings plans from previous meetings were evaluated and planning for the next period was made. Unemployed workers who did not comply could be punished with a benefit sanction, a reduction of UI benefits by 10 percent for a period of two months. In their baseline estimate Van den Berg and Van der Klaauw find no significant treatment effect; that is, counseling and monitoring do not help the unemployed find a job more quickly. However, they also find that counseling and monitoring affect the type of search: unemployed workers who are subject to counseling and monitoring shift their search from informal to formal channels. Van den Berg and Van der Klaauw conclude that focusing monitoring on the unemployed with less favorable characteristics may make more sense because these individuals have less scope for substituting informal for formal search channels. Finally, Graversen and Van Ours (2008) show that mandatory programs help Danish unemployed workers find a job more quickly (see box 12.3 for details).

Box 12.3 ALMPs in Denmark

Brian Krogh Graversen and Jan van Ours analyze data from a Danish experiment that randomly assigned unemployed workers to control and treatment groups on the basis of their birth date. The treatment group was confronted with mandatory activities, whereas the control group was not. The unemployed in the treatment group were informed by letter about their duties within 1 to 2 weeks after becoming unemployed. The letter gave a short description of the activities contained in the program. After 5 to 6 weeks of unemployment individuals had to participate in a job search program that lasted 2 weeks. After the program individuals had to attend meetings once a week or once every second week. The purpose of the meetings was to assist individuals in their job search and to monitor job search efforts. Individuals could also receive job offers mediated by the PES. Before individuals were unemployed for 4 months, they had to receive an offer to participate in an activation program with a duration of at least 3 months. Longer classroom training courses (with a duration of more than 3 months) could not be offered at this stage. Individuals who did not find a job after 6 to 7 months had to participate in a longer meeting with a caseworker, and a new job plan was made. The job plan contained a description of the activities to improve the chances of finding a job. The services offered to the control group during the early stage of the unemployment period were much less intensive than the services offered to the treatment group. Unemployed workers in the control group could voluntarily participate in some

(continued)

Box 12.3 (continued)

of these activities, but this did not happen frequently. Individuals in the control group typically would have to participate in an activation program after one year of unemployment. Graversen and Van Ours find that already before the start of the job search program the job-finding rate in the treatment group was higher than in the control group.

Figure 12.4 presents the survival functions separately for the treatment group and the control group. As shown, the treatment group left unemployment more quickly than the control group. After 3 months 47 percent of the control group and 54 percent of the treatment group had left unemployment. After 6 months 28 percent of the control group was still unemployed, while only 21 percent of the treatment group was still unemployed. The difference between both survival functions increased up to 13 weeks of unemployment, stayed constant until 26 weeks, and declined after that. Figure 12.4 also shows that the median unemployment duration for the control group was about 14 weeks, while for the treatment group it was 11.5 weeks. Clearly there is a substantial treatment effect.

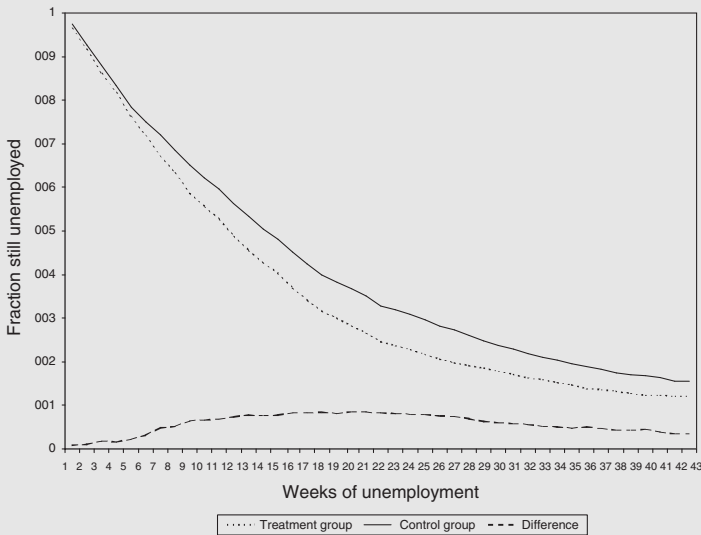


Figure 12.4 Effects of a Danish ALMP Experiment on the Survival Function of Unemployed Workers

Source: Graversen and Van Ours (2008).

12.3.2 Nonexperimental Studies

Nonexperimental studies are based on cross-country comparisons or micro-oriented evaluation studies. Recent evaluation studies, however, are not very optimistic about the benefits of many of these programs. Heckman, Lalonde,

and Smith (1999) provide a detailed overview of microeconomic evaluation studies. They conclude that ALMPs have, at best, a modest impact on participants' labor market prospects. Furthermore, there is considerable heterogeneity in the impact of these programs, so that for some groups of workers the programs are more effective than for other groups of workers. Finally, when programs are implemented on a large scale, displacement and general equilibrium effects may be sizable. This means that if they are not incorporated into a macro framework, micro treatment effect evaluations will provide poor guidance to public policy. Calmfors, Forslund, and Hemström (2001) conclude that the evidence on the effectiveness of Swedish ALMPs is rather disappointing. Labor market retraining, for example, has no or negative employment effects. Martin and Grubb (2001) draw similar conclusions in their overview of what works and what does not work among ALMPs in OECD countries. They conclude that subsidies to employment and direct job creation have not been effective in helping the unemployed get permanent jobs.

If one can draw a general conclusion from the empirical studies based on micro data, it is that the effects of ALMPs on job-finding rates are rather small.⁴ An important drawback of many ALMPs is that they stimulate workers to reduce their search efforts rather than increase them. This is due to the so-called locking-in effect (Van Ours 2004a). Other effects are important too. What is effective for an individual unemployed worker may not be effective in terms of the aggregate level of unemployment. One reason for this may be crowding out. If a training program brings an unemployed worker back to work more quickly at the expense of another unemployed worker finding a job more slowly, the training program is not very efficient. Another reason for the differences between individual and aggregate effects is that a training program may make workers more attractive for firms, which stimulates job creation. It may also be that a training program induces a better match between a worker and a job. In that case job tenure will increase, causing a reduction of unemployment through a reduced inflow into unemployment.

Activation measures seem to be more effective than training programs or employment subsidies. From recent micro studies on the effect of benefit sanctions in the Dutch labor market, we know that a reduction of unemployment benefits may have a substantial effect on the outflow from unemployment to a job. Abbring, Van den Berg, and Van Ours (2005) study the effect of financial incentives by comparing the unemployment duration of individuals who faced a benefit reduction with that of similar individuals who had not been penalized. In the Netherlands people who receive UBs have their benefits reduced if they do not follow the rules related

⁴ Kluge and Schmidt (2002) and Kluge (2006) present overviews of many evaluation studies concerning ALMPs in Europe. They conclude that providing job search assistance and counseling and monitoring accompanied by appropriate sanctions for noncompliance are especially effective, and they are often quite cost-effective because they are rather inexpensive.

to the benefits. According to the UB Law, an unemployed worker has three obligations in order to be entitled to UBs. First, the worker has to prevent unnecessary job loss. Second, the worker has to take actions to prevent staying unemployed (search for a job and accept appropriate job offers, register as a job searcher at the public employment office, participate in education and training, et cetera). Third, the worker has to keep the administrative organization informed about everything that is relevant to the payment of the UB. Related to these obligations are four categories of infringements for which workers can have benefit sanctions imposed: (1) blameworthy unemployment after dismissal, (2) lack of effort to find a job (search intensity too low, declining job offers), (3) administrative infringements (reporting too late), and (4) other infringements (fraud, inaccurate information). The sanction is temporary or permanent (full or partial) reduction of the benefit level. In practice, the temporary partial reduction of the benefits ranged from 5 percent during 4 weeks to 25 or 30 percent during 13 weeks. Abbring, Van den Berg, and Van Ours (2005) analyze how benefit sanctions affect the transition out of unemployment and find that reemployment rates are significantly and substantially raised by imposition of a sanction, from 58 to 67 percent. Van den Berg, Van der Klaauw, and Van Ours (2004) perform a similar analysis on the effects of sanctions on the behavior of Dutch UA recipients in the city of Rotterdam. A benefit sanction raises the transition rate from welfare to work substantially: the job-finding rate more than doubles. A straightforward comparison between individuals who got a benefit sanction imposed and those who did not is incorrect. If this is done, the selectivity of the imposition of a sanction (workers with low motivation who have a low job-finding rate anyway are more likely to get penalized) is not taken into account. Then the estimated sanction effect is seriously underestimated. Although the benefit sanction itself is temporary, the effects turn out to be long-lasting. Even after the sanction period expires, the transition rate from welfare to work is higher than before the sanction was imposed. Similar effects of benefit sanctions are found for the Swiss labor market. From an analysis of Swiss data on benefit sanctions, Lalive, Van Ours, and Zweimüller, (2005) conclude that imposing a benefit sanction reduces unemployment duration by roughly three weeks. They also find evidence of the existence of an ex ante effect.

12.4 Policy Issues

The 1994 OECD *Jobs Study* recommended that governments strengthen the emphasis on active labor market policies and reinforce their effectiveness (OECD 1994). In the reassessment of the Jobs Strategy the OECD concludes “that well-designed programs can have a positive impact . . . but that many existing programs have failed to do so.”⁵ Nevertheless, the OECD concludes that “enough successful

⁵ OECD (2006c), p. 72.

programs have been documented to confirm that an appropriate mix of properly designed ALMPs can reduce unemployment by improving the efficiency of the job-matching process and by enhancing the work experience and skills of those who take part in them.”⁶ This section discusses the main policy issues concerning employment services and activation policies.

12.4.1 Do We Need Public Employment Services?

Employment services are usually provided by government agencies, although there are also private employment services. Public employment services are provided to unemployed workers to assist them in finding jobs. Services range from assistance in locating relevant vacancies and help in job finding to training programs in order to update the skills of the unemployed worker. In order to benefit from market mechanisms, some countries set up quasi-market arrangements in the provision of public employment services and other ALMPs. The OECD (2005b) argues that in order to implement a quasi-market, public employment services have to make a split between a public authority and multiple employment service providers. The public authority is responsible for the determination of individual eligibility for benefits and services, the assigning of the unemployed to specific service providers, and measuring the outcome of the service provision. The local employment service providers can choose their own strategies aimed at bringing their unemployed back to work. The main idea is that through the survival of the fittest, successful strategies will remain even when it is difficult to identify why these strategies are successful. In public employment services the issue of profiling of unemployed workers is important. For an optimal allocation of resources, unemployed workers, with longer expected durations of benefit receipt should receive more or different employment services. There is room for improvement in the functioning of public employment services with regard to profiling of unemployed workers. Lechner and Smith (2007) show that Swiss caseworkers do no better in their profiling than random assignment.

12.4.2 Do We Need Activation Policies?

Activation policies put requirements and obligations on unemployed benefit recipients, such as the obligations to attend intensive interviews with employment counselors, search actively for a new job, and accept job offers. Activation programs differ from public employment services because participation is obligatory for unemployed workers who want to remain entitled to unemployment benefits. Therefore, activation policies make it less attractive for unemployed workers

⁶ OECD (2006b), p. 74.

to collect benefits. This may affect both the inflow of benefit recipients (some unemployed workers will not apply for benefits) and the outflow of benefit recipients (either because the unemployed find jobs or because they drop their benefit claims). Workfare programs provide unemployed workers with temporary jobs in exchange for the payment of unemployment benefits but do not provide additional services. A workfare program may act as a screening device but does not stimulate the unemployed worker to search more intensively. Benefit sanctions imply temporary or permanent reductions of unemployment benefits in order to increase the difference between benefits and postunemployment wages. This will stimulate job search activities. Instead of sticks, like benefit sanctions, carrots, such as wage subsidies, may be used to make it more attractive for workers to search for a job.

12.5 Why Do Active Labor Market Policies Exist?

Active labor market policies exist because unemployment benefits provide disincentives for unemployed workers to find a job. These disincentives may lead to a number of traps (OECD 2005b). The *unemployment trap* arises because benefits paid to unemployed workers and their families are high relative to net income from work. This may discourage job search and put upward pressure on wages (see also chapter 10 on unemployment benefits). The *inactivity trap* is similar to the unemployment trap but concerns individuals of working age who are not receiving unemployment benefits. Often income-related benefits will be lost upon taking paid work. This will reduce the net gains from work. The *poverty trap* is also called the *low-wage trap* because it discourages those already in low-paid work from increasing working hours or moving into higher-paid employment. An increase in gross earnings fails to translate into an increase in net income because of the withdrawal of income-tested benefits. From a cross-national perspective there is a positive correlation between unemployment benefits and ALMPs. Countries with a relatively generous UB system also spend much on ALMP (see also chapter 13). ALMPs stimulate those who are voluntarily unemployed to find a job more quickly than they otherwise would have. However, even without the existence of unemployment benefits, ALMPs can improve the efficiency of the labor market. Unemployed workers who stay unemployed too long may lose some or all of their skills. Involuntarily unemployed workers who want to invest in themselves to acquire the skills necessary to find a job will not always find it easy to finance this investment. Then governments may step in and provide financial support.

12.6 Suggestions for Further Reading

Lars Calmfors (1994) presents a theoretical framework to understand how ALMPs operate. Jim Heckman, Robert LaLonde, and Jeffrey Smith (1999) present a

detailed discussion of the measurement of the effectiveness of ALMPs. Jochen Kluve and Christoph Schmidt (2002) provide a nice overview of many studies exploring the effectiveness of ALMPs. In various years the OECD *Employment Outlook* contains overviews and discussions of ALMP topics.

12.7 Review Questions

1. Why might one criticize subsidized jobs as a policy instrument to bring the unemployed back to work?
2. Why is training possibly more effective in reducing unemployment than subsidized jobs or public employment services?
3. Through what mechanisms do benefit sanctions affect unemployment?
4. How do ALMPs affect the Beveridge curve?
5. In what way does the lump-of-labor fallacy affect the discussion concerning the effectiveness of ALMPs?
6. How do intensive interviews with employment counselors affect the behavior of unemployed workers?
7. What is the relationship between the unemployment trap and upward pressure on wage levels?
8. Why would training be more effective in countries where unemployment benefits are high?
9. The main elements of a system of benefit sanctions are the monitoring intensity and the size of the penalty. Explain the different ways in which these two components affect the behavior of unemployed workers. Use the concepts of ex ante and ex post effects in your explanation.

12.8 Technical Annex: Activating Unemployed Workers

This technical annex shows how a system of benefit sanctions intended to activate unemployed workers can be introduced into a search-matching model of the labor market. In a search-matching model unemployment arises because workers need time to find vacancies and employers need time to find unemployed workers to fill their vacancies. In our model labor is assumed to be homogeneous, and all jobs offer the same wage. The duration of unemployment depends on the intensity of search. The reservation wage is not an instrument of search because that would require heterogeneous labor and a wage distribution. If an unemployed worker and a vacancy match, they produce (nonstorable) output. Individuals and firms are assumed to be risk neutral.

To understand the behavior of workers, the concept of present discounted value (PDV) is used. The PDV of a continuous flow of wages w that lasts forever is

$$V_w = \int_0^{\infty} w e^{-\rho t} dt = \frac{w}{\rho}, \quad (12.1)$$

where ρ is the discount rate. The flow value of having a wage forever the product of the discount rate and the PDV:

$$w = \rho V_w. \quad (12.2)$$

We assume that all workers are homogeneous in the sense that they receive the same wage w net of taxes. Unemployed workers search with intensity s . A worker can adopt search intensity $s \in [0, 1]$ in looking for a job.

Unemployed workers receive unemployment benefits b , where generally $b < w$. In a model without benefit sanctions, the flow value of being unemployed equals the difference between the cost of search and the unemployment benefits and the expected benefits from finding a job

$$\rho V_u = \max_{0 \leq s \leq 1} [b - \gamma(s) + \mu s (V_e - V_u)], \quad (12.3)$$

where μs is the job-finding rate (see technical annex 3.8 for details on other parts of the model).

Activation measures are intended to overcome market failures arising from generous unemployment benefit and welfare benefit schemes. Activation can be introduced through unemployment benefit sanctions, G . We introduce benefit sanctions in a search-matching model (see Boone and van Ours 2006 for details of this model). We make the following assumptions: Each civil servant can monitor ϕ unemployed agents per period; e_g is the number of government officials monitoring the unemployed; u_u is the number of unemployed workers without a benefit sanction; u_s is the number of unemployed workers with a benefit sanction. The Poisson arrival rate of being monitored for an unemployed agent is $\frac{\phi e_g}{u_u}$; because monitoring is imperfect, the probability of being sanctioned, conditional on being monitored, is $(1 - s)$. Then we find that for workers,

$$\rho V_u = \max_{0 \leq s \leq 1} \left[b - \gamma(s) + \mu s (V_e - V_u) + \frac{e_g \phi}{u_u} (1 - s) (V_s - V_u) \right], \quad (12.4)$$

$$\rho V_s = \max_{0 \leq s \leq 1} [(1 - G)b - \gamma(s) + \mu s (V_e - V_s)], \quad (12.5)$$

where we assume that an unemployed worker can receive only one sanction. In this setup the additional term in the value of being unemployed is the term related to the sanctions, a term that reduces the value of unemployment and increases

search intensity. Now, optimal search intensity requires that

$$\gamma'(s_u^*) = \mu (V_e - V_u) - \phi(V_s - V_u), \quad (12.6)$$

$$\gamma'(s_s) = \mu (V_e - V_s). \quad (12.7)$$

We also have to make assumptions concerning the labor market. We assume that government jobs are never destroyed. Therefore, $\rho V_g = w^g$, and in equilibrium $V_e = V_g$. This determines the wage w^g that the government pays its monitoring officials. The labor force $e_g + e + u_u + u_s = 1$. In a steady state labor market the inflow into the stock of unemployed workers without benefit sanctions equals the outflow from that stock. The same holds for the stock of unemployed workers with benefit sanctions. From this we can derive

$$u_u = \frac{\delta^s(1 - e_g)}{\mu s_u + \delta^s} - \frac{e_g \phi(1 - s_u)}{\mu s_s} \frac{\mu s_s + \delta^s}{\mu s_u + \delta^s}, \quad (12.8)$$

$$u_s = \frac{e_g \phi(1 - s_u)}{\mu s_s}.^7 \quad (12.9)$$

The marginal wage tax t is adjusted endogenously in the simulations to satisfy the government budget constraint

$$etw = u_u b + u_s b(1 - G) + e_g w^g, \quad (12.10)$$

where on the right-hand side in the total government expenditures the costs of the benefits system and the monitoring system are included. Benefit sanctions may affect search behavior of unemployed workers in two ways, through the ex ante and ex post effects of benefit sanctions. The ex ante effect refers to the optimal search intensity of workers, which is higher than it would be if workers did not face the possibility of getting a sanction imposed. The ex post effect refers to the effect on the search of having lower benefits once a sanction is imposed. For some systems of benefit sanctions the ex ante effect turns out to be more important than the ex post effect, while for other systems the ex post effect is more important. The comparison between the two depends on the difference between the job arrival rate for workers and the intensity of monitoring. With a low monitoring intensity the possibility of suffering a sanction is small, and the search of the unemployed will not be affected very much. The main effect is after the sanction is imposed. In this case the ex post effect dominates. With a high intensity of monitoring the unemployed will try to reduce the sanction rate by increasing their search intensity. Then the main effect is the ex ante effect. It is even possible that the ex post effect is very small, which in micro research could lead to the erroneous conclusion that

⁷ To see this, note that the equations of motion in the labor market are given by $\dot{u}_u = \delta e - (\mu s_u + (e_g \phi / u_u)(1 - s_u)u_u - \mu s_s u_s)$, and that the steady state is defined as $\dot{u}_u = \dot{u}_s = 0$.

sanctions do not have an effect. They do, but the main effect is in the threat of a penalty, not in the imposition of a penalty.

To illustrate how benefit sanctions affect the labor market, we use the following additional parameter value: monitoring capacity: $\phi = 20$. This parameter value implies that a supervisor deals with 20 unemployed workers and has in a 40-hour week on average about 2 hours per week for every unemployed worker under supervision. This means quite intensive monitoring. The unemployed worker can only avoid benefit sanctions if that person's search intensity is at the maximum value ($s_u = 1$), which is not socially optimal. The results of the simulations are discussed in the main text.

13

Institutional Interactions

Labor market institutions never operate in isolation. Their effects on employment, unemployment, wages, and productivity interact with those of other institutions. Often an institution is introduced just because it counteracts some undesirable effects of another institution. To give an example, the active labor market programs (ALMPs) presented in chapter 12 have been introduced to reduce disincentive effects associated with the provision of unemployment benefits (chapter 11). Sometimes it is optimal to do so because different instruments are needed to achieve different goals. In the preceding example unemployment benefits provide protection against uninsurable labor market risk, while active labor market policies prevent this protection from unduly prolonging unemployment duration by reducing search incentives of UB beneficiaries. Thus the two institutions are complementary. To some extent, ALMPs can also be considered a design feature of unemployment benefit systems.

In other cases institutional complementarities are created as a result of policy failures. For instance, we know from chapter 2 that a minimum wage set at too high a level creates unemployment among low-productivity workers. The solution to this problem is to reduce the statutory minimum wage. However, governments facing strong political resistance to reducing the minimum wages may decide instead to tighten employment protection legislation (chapter 10), reducing the risk of job loss among those low-productivity workers who still have a job. Thus, in order to address a distortion induced by the bad setting of an institution, another distortion is created.

Lack of coordination among different levels of decision making or the actions of uncoordinated pressure groups may also lead to institutional complementarities. For instance, suppose that employment of older workers in a firm is priced out of the market by a collective agreement introducing a steep wage/tenure profile of the type characterized in chapter 6. Without the binding national agreement, the employer and the workers in the plant could agree to introduce some flexibility in wages and working hours among older workers, who would keep their jobs. Because the collective agreement fails to internalize these disemployment effects of the wage-tenure profile, pressure groups operating at the local or plant level may then struggle to force the introduction of early retirement programs. This

type of lack of coordination among different layers of decision making is frequent in the setting of labor market institutions that are not necessarily established by law, but determined within a bargaining process.

Understanding these institutional interactions is important from both a positive and a normative standpoint. From a positive standpoint, the description of the effects of any given institution should take into account all the potential interactions of this institution with other institutional features of the labor market. The difference-in-differences techniques extensively surveyed in this book can identify the effect of any given institution when other institutions are unchanged, but they cannot isolate the effect that, say, EPL has on unemployment when it operates on a stand-alone basis from the effect of EPL on unemployment via minimum wages, unions, and other institutions.

Institutional interactions are important also from a normative standpoint because efficiency-enhancing institutional reforms can find political support only when they are engineered in such a way as to exploit these institutional interactions. To give an example, the strictness of EPL can be reduced by increasing the generosity of unemployment benefit systems along the flexicurity arrangements characterised below in this chapter. Governments need to be aware of any institutional complementarities if they wish to reform labor market institutions. Failure to engineer comprehensive reforms and to take these institutional interactions into account may prevent building a majority in favor of the reforms (Coe and Snower 1997).

The purpose of this chapter is to summarize the key lessons that we have learned in the various chapters about institutional interactions and what empirical observation tells us about the complementarity/substitutability between pairs of institutions. Because the literature on institutional interactions is at its infant stage, the theoretical predictions and especially the empirical findings are far from robust. One should also be aware that the pairwise correlations between institutions displayed in this chapter may well be driven by third factors affecting both institutions, but the presence of these correlations may rule out other explanations and orient further research on institutional interactions.

Because we have been dealing with 11 different institutions in this book, even if we only considered pairs, there would be 110 possible interactions to be considered. Hence we will focus only on the main institutional interactions that (1) are typically observed by empirical work and (2) have been analyzed by labor market theory. As with all the other material in this book, we select issues that we deem both empirically and theoretically more relevant.

13.1 Taxes and Unions

The role of unions in affecting wage setting has been discussed in chapter 3, while the effects of taxation on employment have been dealt with in chapter 4.

The interactions between these two institutions may magnify or reduce their effects on the labor market. We recall briefly the effects of taxes and unions and then consider their potential interactions.

Let us begin with labor taxes. As discussed in chapter 4, when taxes operate in isolation, in a competitive labor market, they do not create unemployment but may reduce employment. The disemployment effects of labor taxes are larger the higher the elasticity of labor supply. When labor supply is rigid, taxes are paid by workers in terms of a reduction in net wages, labor costs for firms are unaffected, and hence there is no effect on employment.

Now consider unions. As discussed in chapter 3, unions modify the effective labor supply faced by individual firms. In particular, unions impose a markup over the reservation wage of individuals, inducing an upward shift of the labor supply faced by employers. We also showed that the level of wage bargaining and the coordination of unions are very important in conditioning wage policies of unions. Centralized wage-setting institutions are more likely to take into account the disemployment effects of wage claims and therefore involve more moderate wage platforms than industry-level or plant-level agreements (Calmfors and Driffil 1988).

Now combine the two institutions. In the presence of unions, labor taxes are bound to affect not only employment but also unemployment because wages are set above the reservation wage of individuals. However, if unions are internally well coordinated or wage setting occurs at the national level, then the adverse effects of taxes on employment are taken into account or internalized in wage setting, inducing unions to reduce wage claims in the presence of labor taxes because they realize that unemployment would increase unless take-home pay concessions are made (Daveri and Tabellini 2000). Encompassing unions, that is, unions engaged in nationwide wage bargaining, would accept accompanying a tax hike with a reduction in net wages, especially if tax proceeds are used for pro-worker policies, including the provision of social transfers to unemployed individuals (Checchi and Lucifora 2002).

The preceding suggests that labor taxes are compatible with low unemployment and wage moderation in widely different circumstances, that is, when unions are weak (as in the United Kingdom since the 1980s) or in the presence of extensive consultation between unions, government, and employers (as, for example, in the Netherlands' recent experience, where welfare reforms were facilitated by wage moderation on the part of corporatist unions; see Nickell and Van Ours 2000). More broadly, in unionized countries labor taxes should have less adverse effects on unemployment when they are coupled with centralized and encompassing collective bargaining institutions. At the same time, taxes on labor are used to finance a variety of social transfers supported by trade unions. Because the trade-off between taxes/transfers and unemployment is more favorable in countries with nationwide wage bargaining and/or strong union

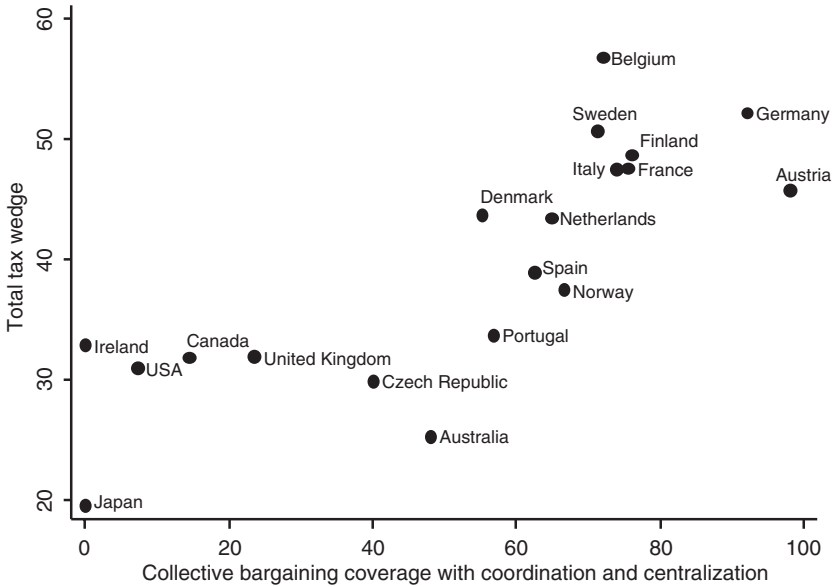


Figure 13.1 Taxes and Unions

Sources: OECD for the indicators and ECHP for the coverage.

Note: The vertical axis displays the total tax wedge measure discussed in chapter 4; the horizontal axis displays the coverage of collective bargaining multiplied by the indexes of centralization/coordination discussed in chapter 3.

coordination (encompassing unions), there may be a complementarity between the two institutions.

Figure 13.1 displays on the vertical axis a measure of the total tax wedge (see chapter 4 for details) and on the horizontal axis the share of workers covered by collective bargaining times an index of centralization/coordination of wage bargaining (the average of the last two columns on the right-hand side of table 3.1, normalized to be in the 0–1 range). The figure points to a positive and statistically significant correlation between the two institutional features: the countries with stronger coverage and coordination/centralization of wage bargaining are those where the tax wedge is larger. This is broadly in line with the preceding theoretical predictions about the interactions between taxes and collective bargaining.

13.2 Employment Protection and Unemployment Benefits

Let us now consider two institutions, unemployment benefits and employment protection, that are supposed to have a similar function: to protect workers against uninsurable labor market risk. As we learned in chapters 10 and 11, EPL and UBs

are only imperfect substitutes. There are three key differences between the two institutions:

1. EPL protects only those who already have a job, acting as a deterrent to layoffs, while UBs protect the population of working age at large (although some work experience is generally required to receive UBs).
2. EPL does not impose any tax burden on workers, while UBs are financed via a payroll tax levied on those who have a job.
3. Under EPL it is the employer having the redundancies who has to offer replacement income to the workers laid off, while UBs are a risk-sharing device that ends- up imposing a fiscal externality on all workers and employers.

These three key differences can be somewhat reduced by appropriate adjustments in the design of UBs or EPL. For instance, UBs can be *experience rated*, as in the United States and, to some extent, in Italy. This means that higher payroll taxes should be paid by those employers who are responsible for the redundancies. This reduces the third difference between EPL and UBs, making UBs closer in their design to a severance payment paid by the former employer. When UBs are experience rated and EPL is a pure *transfer* from the employer to the worker, rather than a *tax*, a payment to a third party (e.g., a lawyer), it does not make much difference for those who have a job whether they are protected by EPL or by UBs. If they are involved in layoffs, they receive, in one case, a mandatory severance payment or, in the other case, a sequence of transfers from the unemployment compensation rolls. Thus the two institutions would be remarkably similar from the standpoint of the individual worker except that the severance provided under EPL is generally provided in a single installment, while UBs offer a stream of (typically monthly) transfers up to the maximum duration of benefits. Also from the standpoint of employers, an experience-rated unemployment benefit system is not much different from severance payments: in both cases they have to pay for their own redundancies.

However, even an experience-rated UB system imposes some fiscal externality on other employers and workers. In order to prevent adverse selection involving only high-turnover employers, UB systems require mandatory payment by all employers. Thus there is always some sharing of the costs of redundancies. Another reason why EPL and UBs cannot be perfect substitutes is that EPL in practice requires that some tax always be imposed on the employers responsible for the dismissals. In order to elicit effort from their workers, employers should establish stronger sanctions for disciplinary than for economic dismissals. In the language of jurisprudence, they should treat differently *subjective just cause* from *objective just cause* dismissals, but this different treatment cannot be generally enforced without envisaging some role for labor judges. Hence EPL, unlike UBs,

always involves some deadweight costs in terms of payment to third parties, like lawyers.

Thus some differences between UBs and EPL are unavoidable, and these differences are perceived by unemployed individuals. As documented in chapter 10, EPL reduces labor turnover, making it more difficult to exit unemployment. UBs are instead more mobility and hiring friendly and, unlike EPL, provide income support during an unemployment spell. Thus unemployed individuals are likely always to rank UBs over EPL. Employees would instead rank EPL over UBs, especially if they have a high discount rate. Political-economic models of labor market institutions (e.g., Wright 1996) predict that configurations with strict EPL arise when the median voter is an insider, a person with a regular contract.

Because of this imperfect substitutability between UBs and EPL, we are unlikely to observe countries with only EPL or only UBs. An additional, political-economic reason for having some of both institutions is that because of the moral hazard problems discussed in chapters 10 and 11, neither EPL nor UBs can provide full insurance against job loss. This means that in countries with more UBs there will be always political pressure to have some EPL, and vice versa. Yet it is also likely that the pressure for one institution will be lower in the countries that have more of the other institution.

The substitutability between UBs and EPL is visible over a cross-section of European countries (figure 13.2): countries with more generous unemployment benefits (the summary measure of benefit entitlement, discussed in chapter 11, times the share of unemployed individuals receiving UBs, or coverage ratio) display less strict EPL, according to the OECD index discussed in chapter 10, and vice versa. The term *flexicurity* is increasingly being used to denote countries, like Denmark, with low EPL and generous UBs. The location of different countries along this trade-off may well reflect the relative political strength of insiders (who support EPL) and outsiders (who support UBs). An example of how changes in the balance of power between these two groups can change the institutional mix along this dimension is provided by Spain, one of the few countries that reduced EPL for regular contracts. Significantly, the reform was enacted in 1994 just after unemployed workers and workers with temporary contracts had become more numerous than employees with regular contracts.

13.3 Unemployment Benefits and Active Labor Market Policies

As discussed in chapter 11, a key problem associated with the provision of UBs relates to disincentives to job search because of imperfect monitoring of job search efforts of unemployment benefit recipients. The active and activation policies characterized in chapter 12 partly reduce these informational asymmetries. In particular, the offer of slots in some active labor market program (subsidized jobs,

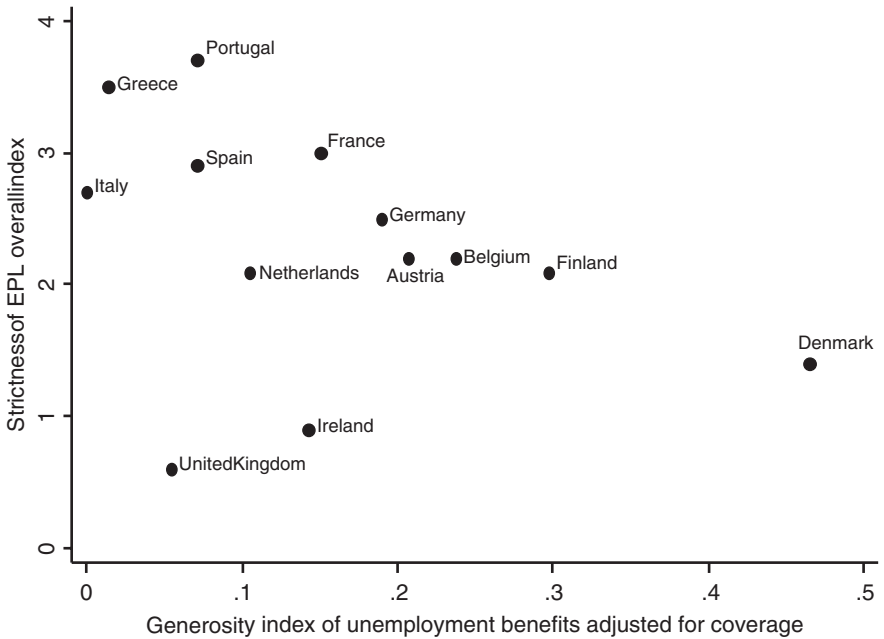


Figure 13.2 Employment Protection and Unemployment Benefits

Sources: OECD for the indicators and ECHP for the coverage.

Note: The vertical axis displays the OECD index of strictness of EPL discussed in chapter 10 and the horizontal axis the OECD summary measure of UB generosity discussed in chapter 11, multiplied by the fraction of unemployed workers receiving the benefits.

training schemes, or public works) can be used as a device to enforce work tests, eliciting whether the UB recipient is actually willing to work and actively seeking employment. Benefit sanctions can also be used as a deterrent to discourage opportunistic behavior by UB recipients.

To the extent that ALMPs reduce moral hazard associated with the provision of UBs, they help reduce the efficiency costs of unemployment insurance, improving the trade-off between insurance and unemployment duration that was characterized in chapter 11. Hence, *ceteris paribus*, one would expect to observe a positive cross-sectional correlation between the generosity of UBs and the investment made by individual countries in policies and infrastructures (e.g., a public employment service) that encourage unemployment benefit recipients to go back to work.

Figure 13.3 displays, on the vertical axis, expenditure on ALMPs as a fraction of GDP and, on the horizontal axis, the summary measure of UB generosity (adjusted by the fraction of unemployed receiving the benefits) that was also used in figure 13.2. The two institutions seem to be positively correlated across countries, consistent with the preceding theoretical insights.

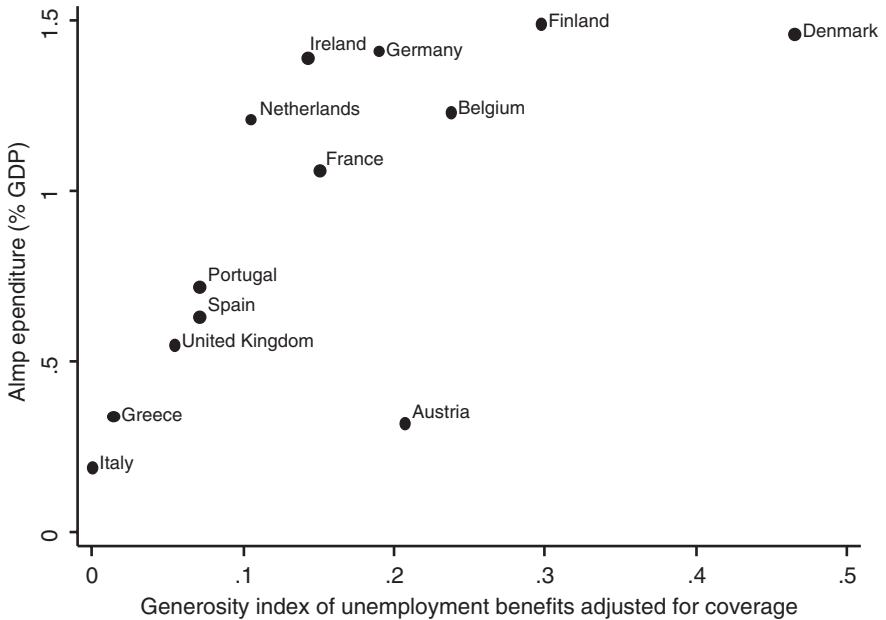


Figure 13.3 Active Labor Market Policies and Unemployment Benefits

Sources: OECD for the indicators and ECHP for the coverage.

Notes: The vertical axis displays expenditure on ALMPs as a fraction of the GDP, as discussed in chapter 12; the horizontal axis displays the OECD summary measure of UB generosity discussed in chapter 11, multiplied by the fraction of unemployed receiving the benefits.

13.4 Employment Protection and Unions

As discussed in chapter 3, unions extract rents from employers by imposing a markup on the reservation wage of workers. They also tend to compress wage structures because they typically pursue egalitarian wage policies. Thus a stronger influence of unions on wage setting is likely to reduce both downward wage adjustment and wage differentials across workers in the same firm, industry, or country.

As pointed out by Bertola and Rogerson (1997), limiting the freedom offered to employers in setting wages induces them to react to shocks by adjusting employment levels. Under compressed wage structures we would expect to observe more turnover in the labor market. By the same token, employment protection would hardly be binding if wages were completely unrestrained over time for a given individual. In response to the labor demand shocks from which EPL is meant to protect workers, wages could fall so as to make stable employment profitable or to induce voluntary quits. Thus there are political-economic reasons to believe that EPL and unions should be complementary institutions.

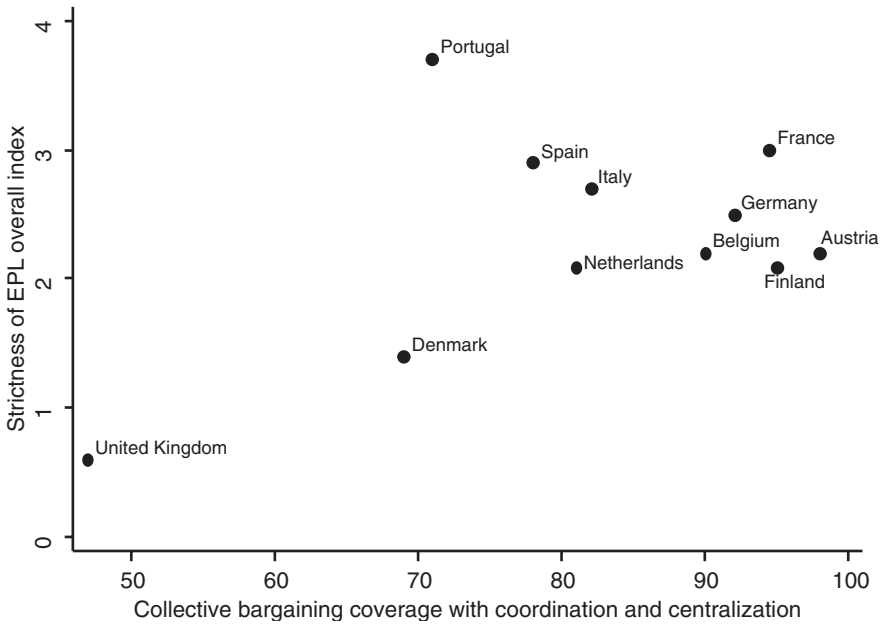


Figure 13.4 Employment Protection and Unions

Sources: OECD for the indicators and ECHP for the coverage.

Note: The vertical axis displays the OECD index of strictness of EPL discussed in chapter 10; the horizontal axis displays the coverage of collective bargaining multiplied by the indexes of centralization/coordination discussed in chapter 3.

Figure 13.4 displays on the vertical axis the measure of collective bargaining coverage and co-ordination/centralization that was also used in figure 13.1. On the horizontal axis we again display the OECD index of strictness of employment protection. The figure points to a positive and statistically significant correlation between the two institutions, in line with the preceding theoretical predictions.

13.5 Taxes and Unemployment Benefits

Taxes on labor are levied to finance various social programs, including UBs. In addition to this *accounting* effect, we may expect to have a positive correlation between UBs and taxes for a number of reasons.

First, economic theory suggests that the effects of labor taxes on employment depends on the use that is made of the tax revenues. Because payroll taxes are often earmarked to unemployment benefit systems, they will be less distortional when they are perceived by workers as a form of insurance. Similar is the case of contributions to public pension systems, which can be perceived by workers as a form of deferred consumption. Under these circumstances workers may be more



Figure 13.5 Taxes and Unemployment Benefits

Sources: OECD for the indicators and ECHP for the coverage.

Note: The vertical axis displays the total tax wedge measure discussed in chapter 4; the horizontal axis displays the OECD summary measure of UB generosity discussed in chapter 11, multiplied by the fraction of unemployed workers receiving the benefits.

keen to accept reductions of their current net wage because they know that these reductions will improve their living standards later in their life.

Second, as discussed in chapter 11, contributions to unemployment insurance are typically proportional to the (gross) wage bill, but there are floors and ceilings in the provision of UBs. Hence even if all workers (and employers) pay the same fraction of their gross wages for UBs, the funding of UBs is *de facto* progressive. Unemployment assistance (the income support schemes offered in long durations of unemployment) are typically funded via general government revenues, drawing on taxes, for example, income taxes, which are typically progressive. As pointed out in chapter 4, taxes have less effect on employment when they are progressive. Progressive taxation, by reducing labor taxes at low earning levels, reduces the adverse effects of taxes on labor supply and unemployment. Hence to the extent that UB financing involves an important degree of redistribution (see chapter 11), more UBs can better coexist with high labor taxes.

Figure 13.5 displays the correlation between the tax wedge on labor and the measure of the generosity of UBs already used in this chapter (e.g., in figure 13.3). The figure hints at a positive correlation between the two institutional features: countries with more taxes on labor provide more generous unemployment benefits.

13.6 Education and Retirement Programs

Both education and retirement programs operate intergenerational redistributions, albeit typically in opposite directions. Education policies, increasing the quality and availability of education, are bound to improve the welfare of younger generations, generally by taxing current workers. Through the public financing of education, the young borrow from the middle-aged to invest in human capital. Generous pay-as-you-go retirement plans instead tend to improve the well-being of older workers at the expense of younger generations who have to pay for their pensions.

The degree of representation of the different generations in the political process is therefore bound to affect the size and design of these two institutions. In general, the older the population and the stronger the representation of the interests of the older generations, the more generous should be the pay-as-you-go system and the lower the expenditure on education. Thus political-economic equilibriums should in principle imply an inverse cross-country correlation between generosity of public pension systems and public expenditure on education. In general, aging societies should experience a shift in the policy mix in favor of more generous retirement institutions (Galasso 2006) and against large-scale public education programs. However, the two institutions also provide the (missing) market for intergenerational transfers. Young people borrow from middle-aged people in order to invest in their human capital, and they pay back their debt to the previous generation through of social security contributions. Under this interpretation, the two institutions could only jointly improve the efficiency of intergenerational transfers, making retirement schemes and education programs support each other (Boldrin and Montes 2005).

Furthermore, the increase in longevity also involves an increase in the time horizon considered when decisions to invest in human capital are made. The longer the expected duration of working life, the stronger the incentives to make an initial investment in education in order to obtain higher earnings while working. It is the same principle that makes migration largely concentrated in young people: they have more time to capitalize on their initial investment (chapter 9). This institutional interaction, therefore, should induce a positive correlation between the two institutions.

Education was also found in chapter 6 to improve productivity at older ages. This effect militates in favor of a substitutability between the two institutions because people who are highly productive at later stages of their working life are likely to work longer and need less generous retirement plans. However, as argued earlier, education policies target different generations than those involved in contemporaneous retirement plans. Thus there are theoretical arguments militating in both directions, toward either complementarity or substitutability between the two institutions.

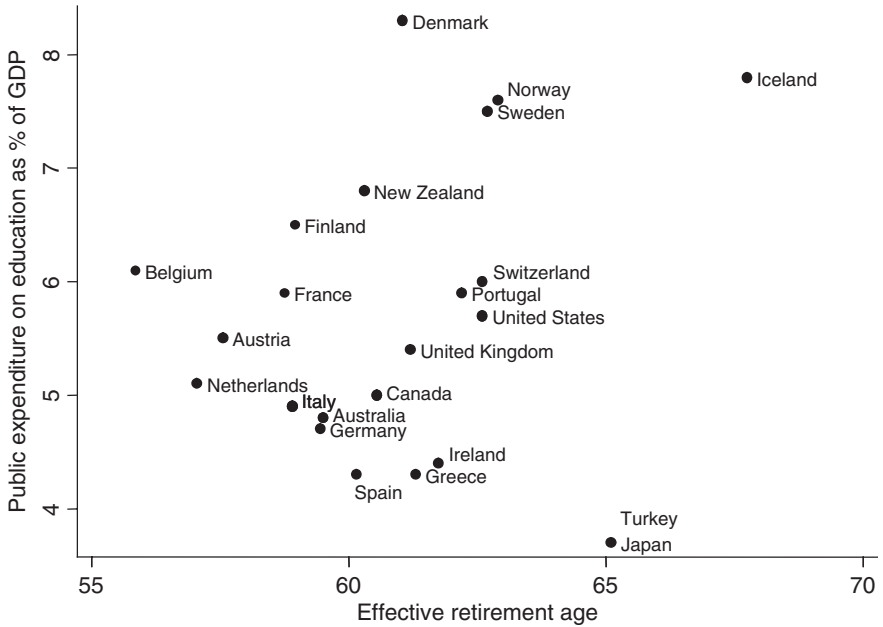


Figure 13.6 Public Education and Early Retirement

Source: OECD.

Note: The vertical axis displays expenditure on public education as a fraction of GDP according to the definitions discussed in chapter 8; the horizontal axis measures the effective age of retirement as discussed in chapter 6.

With these caveats in mind, figure 13.6 displays on the vertical axis the share of public education expenditure in GDP and on the horizontal axis the effective retirement age in several OECD countries. There is not a clear pattern of correlation between the two institutions. Nordic countries typically display relatively long working lives and high expenditure on education, while southern European countries feature short working life and low expenditure on education. However, there are also countries like Belgium, with high education expenditure and early retirement, or Japan, with relatively long working lives but low spending on education.

13.7 What Future for These Institutions?

As we have seen, some institutions seem to be correlated across countries and in a way that is generally consistent with predictions of economic theory. Much more work is needed to understand the underlying transmission mechanisms. Institutional interactions are a fascinating research agenda whose surface has only been scratched here. It is important to keep in mind that the pairwise correlations highlighted in this chapter do not tell us anything about the order of causation:

institution A may be affecting institution B, or it may be just the other way around. Moreover, we observed that the same institution can be correlated with several institutions at the same time. For instance, we found labor taxes to be positively correlated not only with (encompassing) unions but also with unemployment benefits. This suggests that unions and unemployment benefits are also positively correlated, and we can possibly find theoretical explanations for a correlation between unions and unemployment benefits as well. As this example makes clear, the pairwise correlations displayed in this chapter may well capture the effect of third factors, affecting several institutions at the same time and driving the correlations.

One of these factors is social customs. There is an increasing body of literature on the cultural determinants of labor market institutions (Algan and Cahuc 2006). Civic attitudes, in particular, can significantly reduce moral hazard problems related to the provision of unemployment benefits and increase the attractiveness of this institution versus, for instance, employment protection legislation, favoring Danish-type flexicurity arrangements with low EPL and high UBs (Boeri, Galasso, and Conde-Ruiz 2007). More broadly, strong social sanctions against opportunistic behavior may significantly increase the efficiency of social transfers and reduce the need of employment and wage rigidities.

Another key common factor that is likely to drive institutional correlations is the degree of competition in product markets. Larger rents are likely to induce more political support for employment protection (Saint-Paul 2000) because they increase the value of holding a job versus unemployment status. Higher rents also create stronger support for unions to extract rents from employers: as shown in chapter 3, unions are stronger in industries where there is less competition. Thus product market rents may be responsible for the positive correlation between EPL and unions that was documented in this chapter and for other correlations as well.

Insofar as culture and product market rents drive these institutional correlations, we may expect globalization to reduce institutional complementarities. As discussed in chapter 1, globalization makes labor demand more elastic, reducing cross-country asymmetries in the extent of product market rents. Globalization also redefines the borders between tradable and nontradable goods, reducing rents also in sectors traditionally sheltered from competition in international markets. Globalization may also induce a reduction in cultural differences. This is a longer-term process and one that is far from being understood. It is also possible that local communities may react to globalization by increasing their cultural differences with respect to the rest of the world.

Other phenomena shaping labor market institutions are occurring on a global scale, although at a different speed. In particular, the womenization of the workforce and the aging of populations are global phenomena. These developments are likely to give more prominence in the institutional mix to policies reconciling work and family life and allowing for more flexibility throughout the life course.

Because human capital investment is essential in preserving productivity at higher ages, it is also likely that greater attention will be devoted to formal education, as well as lifelong training.

To the extent that cross-country asymmetries in the common factors driving institutional interactions are reduced, we are bound to observe a reduced importance of national clusters of institutions. An extensive literature on national labor market and social policy models (Bertola et al. 2001) provides undisputed taxonomies of countries (the Scandinavian model, the continental European model, the Anglo-Saxon model, the Southern model, and so on). However, these taxonomies are becoming less and less powerful: the within-model variance in the levels of different institutions is increasing relative to the dispersion across country groupings. Thus there is more mixing of the models rather than less heterogeneity across countries. The fact that cross-country differences in the scale and design of labor market institutions are not fading away can also be explained by the fact that labor market institutional design is an ongoing field of experimentation. Different countries are trying in different ways to reconcile protection against uninsurable labor market risk with better work incentives. There is much to learn from these ongoing experiments. Research in labor economics can greatly contribute to a steeper learning process. It is essential that researchers look very carefully at institutional design and implementation. A key message that we hope that this book has succeeded in conveying is that the devil is in the details.

References

- Abbring, J. H., G. J. van den Berg, and J. C. van Ours. 2005. The effect of unemployment insurance sanctions on the transition rate from unemployment to employment. *Economic Journal* 115:602–30.
- Abowd, J. M., and F. Kramarz. 2000. Inter-industry and firm-side wage differentials: New evidence from limned employer-employee data. Ithaca, NY: Cornell University mimeo, unpublished.
- Abowd, J., F. Kramarz, T. Lemieux, and D. Margolis. 1999. Minimum wage and youth employment in France and the United States. In *Youth employment and the labor market*, ed. D. Blanchflower and R. Freeman. Chicago: University of Chicago Press.
- Acemoglu, D. 1997. Training and innovation in an imperfect labor market. *Review of Economic Studies* 64:445–64.
- . 2001. Good jobs versus bad jobs: Theory and some evidence. *Journal of Labor Economics* 19:1–22.
- Acemoglu, D., P. Aghion, and G. Violante. 2000. Deunionization, technical change, and inequality. Carnegie-Rochester Conference Series on Public Policy.
- Acemoglu, D., and J. Pischke. 1998. Why do firms train? Theory and evidence. *Quarterly Journal of Economics* 113:79–119.
- . 1999. Beyond Becker: Training in imperfect labor markets. *Economic Journal* 109:F112–F142.
- Acemoglu, D., and R. Shimer. 1999. Efficient unemployment insurance. *Journal of Political Economy* 107:893–928.
- . 2000. Wage and technology dispersion. *Review of Economic Studies* 67:585–608.
- Adams, S. J. 2004. Age discrimination legislation and the employment of older workers. *Labor Economics* 11:219–41.
- Addison, J., and M. Blackburn. 1999. Minimum wages and poverty. *Industrial and Labor Relations Review* 52 (3): 393–409.
- Addison, J., and J.-L. Grosso. 1996. Job security provisions and employment: Revised estimates. *Industrial Relations* 35 (4): 585–603.
- Agell, J. 2000. On the determinants of labor market institutions: Rent-sharing vs. social insurance. Working Papers Uppsala, Working Paper Series no. 16.
- Agell, J., and K. E. Lommerud. 1992. Union egalitarianism as income insurance. *Economica* 59:295–310.
- Akerlof, G. A. 1970. The market for ‘lemons’: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84 (3):488–500.
- . 1980. A theory of social customs, of which unemployment may be one consequence. *Quarterly Journal of Economics* 94 (4): 749–75.

- Akerlof, G. A., and J. Yellen, eds. 1986. *Efficiency wage theories of the labor market*. Cambridge: Cambridge University Press.
- Alesina, A., and R. Perotti. 1997. The welfare state and competitiveness. *American Economic Review* 87 (5): 921–39.
- Algan, Y., and P. Cahuc. 2005. The roots of low european employment: Family culture? In *NBER macroeconomics annual*, ed. C. Pissarides and J. Frenkel. Cambridge, MA: MIT Press.
- . 2006. Civic attitudes and the design of labor market institutions: Which countries can implement the Danish flexicurity model? CEPR Discussion Papers no. 5489, London.
- Alogoskoufis, G., C. R. Bean, G. Bertola, D. Cohen, J. J. Dolado, and G. Saint-Paul. 1995. *Unemployment: Choices for Europe*. Monitoring European integration 5. CEPR Report. London: CEPR.
- Alvarez-Plata, P., H. Brücker, and B. Siliverstovs. 2003. *Potential migration from central and Eastern Europe into the EU-15—An update*. Report for the European Commission. Brussels: DG Employment and Social Affairs.
- Ames, E., and R. T. Rapp. 1977. The birth and death of taxes: A hypothesis. *Journal of Economic History* 37:161–78.
- Andrews, M. J., T. Schank, and R. Simmons. 2005. Does worksharing work? Some empirical evidence from the IAB-establishment panel. *Scottish Journal of Political Economy* 52: 141–76.
- Angrist, J. D. 1990. Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records. *American Economic Review* 80 (3): 313–36.
- . 1991. Instrumental variables estimation of average treatment effects in econometrics and epidemiology. NBER Technical Working Papers 0115, National Bureau of Economic Research, Cambridge, MA.
- . 1999. Empirical strategies in labor economics. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 1277–1366, Amsterdam: Elsevier.
- Angrist, J. D., and A. B. Krueger. 1991. Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics* 106 (4): 979–1014.
- . 1999. Empirical strategies in labor economics. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 1277–1366 Amsterdam: Elsevier.
- Antecol, H., D. Cobb-Clark, and S. Trejo. 2003. Immigration policy and the skills of immigrants to Australia, Canada and the United States. *Journal of Human Resources* 38 (1): 192–218.
- Ashenfelter, O. C., and J. N. Brown. 1986. Testing the efficiency of employment contracts. *Journal of Political Economy* 94 (3): 541–87.
- Ashenfelter, O. C., and D. Card. 2002. Did the elimination of mandatory retirement affect faculty retirement flows? *American Economic Review* 92:957–80.
- Ashenfelter, O. C., and C. Rouse. 1998. Income, schooling, and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics* 113:253–84.
- Atkinson, A., and J. Micklewright. 1991. Unemployment compensation and labor market transitions: A critical review. *Journal of Economic Literature* 29:1679–1727.
- Aubert, P. 2003. Productivity, wage and demand for elder workers: An examination of French matched employer-employee data. Working Paper, INSEE, Paris.
- Autor, D., J. Donohue, and S. Schwab. 2006. The costs of wrongful discharge laws. *Review of economics and statistics* 88 (2): 211–31.

- Avolio, B. J., D. A. Waldman, and M. A. McDaniel. 1990. Age and work performance in non-managerial jobs: The effects of experience and occupational type. *Academy of Management Journal* 33 (2): 407–22.
- Aydemir, A., and G. Borjas. 2006. A comparative analysis of the labor market impact of international migration: Canada, Mexico and the United States. NBER Working Paper no. W1232.
- Bacache-Beauvallet, M., and E. Lehmann. 2005. Minimum wage or negative income tax: Why skilled workers may favor wage rigidities. IZA Discussion Paper, no. 1570.
- Baldwin, R. E., J. F. Francois, and R. Portes. 1997. The costs and benefits of eastern enlargement: The impact on EU and central Europe. *Economic Policy* 24:97–105.
- Banks, J., and S. Smith. 2006. Retirement in the UK. *Oxford Review of Economic Policy* (Spring): 40–56.
- Barro, R., and X. Sala-i-Martin. 1991. Convergence across states and regions. *Brookings papers on economic activity* 1:107–82.
- . 1995. *Economic growth*. New York: McGraw-Hill.
- Barth, M. C., W. McNaught, and P. Rizzi. 1993. Corporations and the aging workforce. In *Building the competitive workforce: Investing in human capital for corporate success*, ed. P. H. Mirvis. New York: John Wiley and Sons.
- Bassanini, A., and R. Duval. 2006. Employment patterns in OECD countries: Reassessing the role of policies and institutions. OECD Economics Department Working Papers no. 486, Paris.
- Bauer, T. K., and K. F. Zimmermann. 1997. Looking south and east: Labor market implications of migration in Europe and LDCs. In *Globalisation and labor markets: Challenges, adjustment and policy responses in the EU and the LDCs*, ed. O. Memedovic, A. Kuyvenhoven, and W.T.M. Molle, 75–103. Dordrecht: Kluwer.
- . 1999. Assessment of possible migration pressure and its labor market impact following EU enlargement to central and Eastern Europe. IZA Research Report no. 3, Bonn.
- Bean, C. 1994. European unemployment: A survey. *Journal of Economic Literature* 32 (2): 573–619.
- Bean, C., R. Layard, and S. Nickell. 1986. The rise in unemployment: A multi-country setting. *Economica* 53:S1–S22.
- Becker, G. S. 1964. *Human capital: A theoretical and empirical analysis, with special reference to education*. New York: Columbia University Press.
- Behaghel, L., B. Crépon, and B. Sédillot. 2004. Contribution Delalande et transitions sur le marché du travail. *Economie et Statistique* 372:61–88.
- Bell, L. A. 1997. The impact of minimum wages in Mexico and Columbia. *Journal of Labor Economics* 15:S102–S135.
- Belot, M., and J. van Ours. 2001. Unemployment and labor market institutions: An empirical analysis. *Journal of the Japanese and International Economies* 15:403–18.
- Benhabib, J. 1996. On the political economy of immigration. *European Economic Review* 40 (9):1737–43.
- Bentolila, S., and J. Dolado. 1994. Labor flexibility and wages: Lessons from Spain. *Economic Policy* 18:55–99.
- Bertola, G. 1990. Job security employment and wages. *European Economic Review*, 851–66.
- Bertola, G., F. Blau, and L. Kahn. 2002. Labor market institutions and demographic employment. NBER Working Paper no. 9043, Cambridge, MA.

- Bertola, G., and T. Boeri. 2002. EMU labour markets two years on: Microeconomic tensions and institutional evolutions. In *EMU and economic policy in Europe*, ed. M. Buti and A. Sapir. Cheltenham, UK and Northampton, MA: Edward Elgar.
- Bertola, G., T. Boeri, and S. Cazes. 2000. Employment protection in industrialized countries: The case for new indicators. *International Labor Review* 139 (1): 57–72.
- Bertola, G., J. F. Jimeno, R. Marimon, and C. Pissarides. 1999. EU welfare systems and labor markets: Diverse in the past, integrated in the future? In *Welfare and employment in Europe*, ed. G. Bertola, T. Boeri, and G. Nicoletti. 2001. Cambridge, MA: MIT Press.
- Bertola, G., and R. Rogerson. 1997. Institutions and labour reallocation. *European Economic Review* 41:1147–71.
- Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. The Nash solution in economic modelling. *Rand Journal of Economics* 17 (2): 176–88.
- Björklund, A., and B. Holmlund. 1991. The economics of unemployment insurance: The case of Sweden. In *Labour market policy and unemployment insurance*, ed. A. Björklund, R. Haveman, R. Hollister, and B. Holmlund. Oxford: Clarendon Press.
- Black, D. A., J. A. Smith, M. C. Berger, and B. J. Noel. 2003. Is the threat of training more effective than training itself? Evidence from random assignments in the UI system. *American Economic Review* 93 (4): 1313–27.
- Blanchard, O. 2006. European unemployment: The evolution of facts and ideas. *Economic Policy* 22:5–59.
- Blanchard, O., and L. H. Katz. 1992. Regional evolutions. *Brookings Papers on Economic Activity*, 1–75.
- Blanchard, O., and J. Tirole. 2003. Redesigning the employment protection system. *De Economist* 152:1–20.
- Blanchard, O., and J. Wolfers. 2000. The role of shocks and institutions in the rise of European unemployment: The aggregate evidence. *Economic Journal* 110:1–33.
- Blanchard, P., and O. Blanchard. 1989. The Beveridge curve. *Brookings Papers on Economic Activity* 1:1–76.
- Blanchflower, D. 2007. A cross-country study of union membership. *British Journal of Industrial Relations* 45 (1): 1–28.
- Blau, D. M. 1992. The child care labor market. *Journal of Human Resources* 27 (1): 9–39.
- . 2000. Child care subsidy programs. NBER Working Paper no. 7806, Cambridge, MA.
- . 2001. *The child care problem: An economic analysis*. New York: Russell Sage Foundation.
- . 2003. Reconciling work and family. CESifo DICE Report 4:3–7.
- Blau, F. D., and L. M. Kahn. 1996. The gender earnings gap: Some international evidence. NBER Working Paper no. 4224, Cambridge, MA.
- Blöndal S., and S. Scarpetta. 1998. The retirement decision in OECD countries. OECD Economics Department Working Paper no. 202, Paris.
- Blonigen, B. A., R. B. Davies, and K. Head. 2003. Estimating the knowledge-capital model of the multinational enterprise: Comment. *American Economic Review* 93: 980–94.
- Blundell, R. 2001. Welfare reform for low income workers. *Oxford Economic Papers* 53, 189–214.
- Blundell, R., and M. Dias. 2000. Evaluation methods for non-experimental data. *Fiscal Studies* 21:427–68.

- Blundell, R., and T. MaCurdy. 1999. Labor supply: A review of alternative approaches. In *Handbook of labor economics*, ed. O. C. Ashenfelter and D. Card, 1559–1695. Amsterdam: Elsevier.
- Böckerman, P., and R. Uusitalo. 2006. Erosion of the Ghent system and union membership decline: Lessons from Finland. *British Journal of Industrial Relations* 44 (2): 283–303.
- Boeri, T. 1996. Is job turnover countercyclical? *Journal of Labor Economics* 14 (4): 603–25.
- . 1999. Enforcement of employment security regulations, on-the-job search and unemployment duration. *European Economic Review* 43:65–89.
- . 2000. *Structural change, welfare systems and labor reallocation*. Oxford: Oxford University Press.
- . 2005. An activating social security system. *De Economist* 153 (4): 375–97.
- . 2008. Migration policy and the welfare state. IZA Discussion Paper, Bonn.
- Boeri, T., and H. Bruecker. 2005. Why are Europeans so tough on migrants? *Economic Policy* 44:629–704.
- Boeri, T., A. Brugiavini, and L. Calmfors, eds. 2001. *The role of unions in the twenty-first century*. Oxford: Oxford University Press.
- Boeri, T., and M. C. Burda. 2008. Preferences for rigid versus individualized wage setting in search economies with firing frictions. Discussion Paper.
- Boeri, T., M. C. Burda, and F. Kramarz. 2008. *Are Europeans lazy? Or Americans crazy?* Oxford University Press, forthcoming.
- Boeri, T., I. Conde-Ruiz, and V. Galasso. 2007. The political economy of flexicurity. FEDEA Working Paper no. 2006–15, Madrid.
- Boeri, T., D. Del Boca, and C. Pissarides, eds. 2005. *Women at work: An economic perspective*. Oxford: Oxford University Press.
- Boeri, T., and P. Garibaldi. 2006. Two tier reforms of employment protection: A honeymoon effect? *Economic Journal* 117:357–85.
- Boeri, T., G. Hanson, and B. McCormick, eds. 2002. *Immigration policy and the welfare system*. Oxford: Oxford University Press.
- Boeri, T., and J. Jimeno. 2005. The effects of employment protection: Learning from variable enforcement. *European Economic Review* 49 (8): 2057–77.
- Boeri, T., R. Layard, and S. Nickell. 2001. Welfare-to-work and the fight against long-term unemployment. Department for Education and Employment, Research Report U.K. no. 206.
- Boeri, T., and M. Macis. 2008. Do unemployment benefits promote or hinder structural change? Unpublished manuscript.
- Boeri, T., and G. Tabellini. 2001. Does information increase political support for pension reform? Mimeo.
- Boldrin, M., J. J. Dolado, J. F. Jimeno, and F. Peracchi. 1999. The future of pension systems in Europe: A reappraisal. *Economic Policy* 29:289–323.
- Boldrin, M., and A. Montes. 2005. The intergenerational state education and pensions. *Review of Economic Studies* 72:651–64.
- Bonin, H., B. Raffelhüschen, and J. Walliser. 2000. Can immigration alleviate the demographic burden? *FinanzArchiv* 57:1–21.
- Boone, J., and J. C. van Ours. 2004. Effective active labor market policies. Discussion Paper 04-87, Center, University of Tilburg.

- . 2006. Modelling financial incentives to get unemployed back to work. *Journal of Institutional and Theoretical Economics* 162 (2):227–52.
- Booth, A. L. 1995. *The economics of the trade union*. Cambridge: Cambridge University Press.
- Booth, A. L., and M. Bryan. 2005. Testing some predictions of human capital theory: New training evidence from Britain. *Review of Economics and Statistics* 87:391–94.
- Booth, A. L., and M. Chatterji. 1995. Union membership and wage bargaining when membership is not compulsory. *Economic Journal* (Royal Economic Society) 105 (429): 345–60.
- Booth, A. L., and M. Ravallion. 1993. Employment and length of the working week in a unionized economy in which hours of work influence productivity. *Economic Record* (Economic Society of Australia) 69 (207): 428–36.
- Borjas, G. J. 1987. Self-selection and the earnings of immigrants. *American Economic Review* 77 (4): 531–53.
- . 1994. The economics of immigration. *Journal of Economic Literature* 32: 1668–1717.
- . 1999. The economic analysis of immigration. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 3:1679–1760. Amsterdam: Elsevier.
- . 2001. Does immigration grease the wheels of the labor market? *Brookings Papers on Economic Activity* 1:69–133.
- . 2003. The labor demand curve is downward-sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics* 118 (November): 1335–74.
- . 2006. *Labor economics*. 4th ed. Cambridge: McGraw Hill.
- Brandolini, A., P. Cipollone, and E. Viviano. 2004. Does the ILO definition capture all unemployment? *Journal of the European Economic Association* 4:153–70.
- Bratberg, E., T. H. Holmås, and Ø. Thøgersen. 2004. Assessing the effects of an early retirement program. *Journal of Population Economics* 17:387–408.
- Brown, C. 1999. Minimum wages, employment, and the distribution of income. In *Handbook of labor economics*, ed. O. Ashenfelter, and D. Card, 3B:2102–63. Amsterdam: Elsevier.
- Brücker, H. 2001. Die Folgen der Freizügigkeit für die Ost-West-Migration: Schlussfolgerungen aus einer Zeitreihenanalyse der Migration nach Deutschland. *Konjunkturpolitik—Applied Economics Quarterly*, Supplement 52:17–54.
- . 2002. Can international migration solve the problems of European labor markets? *UNECE Economic Survey of Europe* 2:109–42.
- Brücker, H., and M. Kohlhaas. 2004. Migration, qualification and labor markets. Mimeo, German Institute for Economic Research (DIW), Berlin.
- Brücker, H., and P. J. H. Schröder. 2005. International migration with heterogeneous agents. Mimeo, Aarhus School of Business.
- Brücker, H., and B. Siliverstovs. 2004. On the estimation and forecasting of international migration: How relevant is heterogeneity across countries? DIW Discussion Paper no. 314, Berlin.
- Brugiavini, A., M. C. Burda, L. Calmfors, D. Checchi, R. Naylor, and J. Visser. 2001. The future of collective bargaining in Europe. In *The role of unions in the twenty-first century*, ed. T. Boeri, A. Brugiavini, and L. Calmfors. Oxford: Oxford University Press.

- Brunello, G., P. Garibaldi, and E. Wasmer. 2007. *Education and training in Europe*. Oxford: Oxford University Press.
- Buddelmeyer, H., G. Mourre, and M. Ward. 2005. Part-time work in EU Countries: Labour market mobility, entry and exit. IZA Discussion Paper 1550, Bonn.
- Burda, M. C. 1995. Migration and the option value of waiting. *Economic and Social Review* 27 (1): 1–19.
- Burda, M. C., and J. Hunt. 2001. From reunification to economic integration: productivity and the labor market in eastern Germany. *Brookings Papers on Economic Activity* 2:1–92.
- Burda, M. C., and P. Weil. 2006. Blue laws. Unpublished paper, Humboldt-Universität zu Berlin and Université Libre de Bruxelles.
- Burdett, K., and D. Mortensen. 1998. Wage differentials, employer size, and unemployment. *International Economic Review* 30:257–73.
- Burtless, G. 1985. Are targeted wage subsidies harmful? Evidence from a wage voucher experiment. *Industrial and Labor Relations Review* 39:105–14.
- Cahuc, P., and J. Michel. 1996. Minimum wage, unemployment and growth. *European Economic Review* 40:1463–82.
- Cahuc, P., and A. Zylberberg. 2004. *Labor economics*. Cambridge, MA: MIT Press.
- Calmfors, L. 1994. Active labor market policy and unemployment—A framework for the analysis of crucial design features. *OECD Economic Studies* 22 (1): 7–47.
- . 1995. What can we expect from active labor market policy? *Konjunkturpolitik* 43:11–30.
- Calmfors, L., and J. Driffil. 1988. Bargaining structure, corporatism, and macroeconomic performance. *Economic Policy* 6:12–61.
- Calmfors, L., A. Forslund, and M. Hemström. 2001. Does active labor market policy work? Lessons from the Swedish experiences. *Swedish Economic Policy Review* 85:61–124.
- Calmfors, L., and M. Hoel. 1988. Work sharing and overtime. *Scandinavian Journal of Economics* 90:45–62.
- . 1989. Work sharing, employment and shiftwork. *Oxford Economic Papers* 41:758–73.
- Card, D. 1990. The impact of the Mariel Boatlift on the Miami labor market. *Industrial Labor Relations Review* 43:245–57.
- . 1992. The effects of unions on the distribution of wages: Redistribution or relabelling. NBER Working Paper no. 4195, Cambridge, MA.
- . 1999. The causal effect of education on earnings. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 1801–63. Amsterdam: Elsevier.
- . 2001. The effects of unions on wage inequality in the U.S. labor market. *Industrial and Labor Relations Review* 54:296–315.
- . 2005. Is the new immigration really so bad? *Economic Journal* 115(507): 300–324.
- Card, D., and D. R. Hyslop. 2005. Estimating the effects of a time-limited earnings subsidy for welfare-leavers. *Econometrica* 73:1723–70.
- Card, D., and A. Krueger. 1994. Minimum wage, and employment: A case study of the fast food industry in New Jersey and Pennsylvania. *American Economic Review* 84 (4): 772–93.
- . 1995. *Myth and measurement: The new economics of the minimum wage*. Princeton, NJ: Princeton University Press.
- . 2000. Minimum wages: A case study of the fast food industry in New Jersey and Pennsylvania: Reply. *American Economic Review* 90 (5): 1397–1420.

- Card, D., and P. Levine. 2000. Extended benefits and the duration of UI spells: Evidence from the New Jersey Extended Benefit Program. *Journal of Public Economics* 78:107–38.
- Carone, G., D. Costello, N. Diez Guardia, G. Mourre, B. Przyware, and A. Salomaki. 2005. The economic impact of aging populations in the EU25 Member states. Economic Papers no. 236, European Commission, Brussels.
- Carr, D., J. R. Markusen, and K. E. Maskus. 2001. Estimating the knowledge-capital model of the multinational enterprise. *American Economic Review* 91:693–708.
- Carrington, W., and P. de Lima. 1996. The impact of the 1970s repatriates from Africa on the Portuguese labor market. *Industrial and Labor Relations Review* 49 (January):330–47.
- Chan, S., and A. H. Stevens. 1999. Employment and retirement following a late-career job loss. *American Economic Review* 89:211–16.
- . 2001. Job loss and employment patterns of older workers. *Journal of Labor Economics* 19:484–521.
- Checchi, D., and C. Lucifora. 2002. Unions and labor market institutions in Europe. *Economic Policy* 35:363–408.
- Coe, D., and D. Snower. 1997. Policy complementarities: The case for fundamental labor market reforms. CEPR Discussion Papers no. 1585, London.
- Collins, W. J., K. O'Rourke, and J. G. Williamson, 1999. Were trade and factor mobility substitutes in history? In *Migration: The controversies and the evidence*, ed. R. Faini, J. De Melo, and K. F. Zimmermann, 227–59. Cambridge: Cambridge University Press.
- Contensou, F., and R. Vranceanu. 2000. *Working time: Theory and policy implications*. Cheltenham, UK and Northampton, MA: Edward Elgar.
- Corneo, G. 1995a. Distributional implications of a shorter working week: An unpleasant note. *Journal of Economics* 62:25–31.
- . 1995b. National wage bargaining in an internationally integrated product market. *European Journal of Political Economy* 11:107–16.
- . 1995c. Social custom, management opposition and trade union membership. *European Economic Review* 39:275–92.
- . 1997a. Taxpayer-consumers and public pricing. *Economics Letters* 57:235–40.
- . 1997b. The theory of the open shop trade union reconsidered. *Labor Economics* 4:71–84.
- Costa, L., and M. Kahn. 2000. Power couples: Changes in the location choice of the college educated. *Quarterly Journal of Economics* 115:1287–1314.
- Crépon, B., N. Deniau, and S. Pérez-Duarte. 2003. Wages, productivity, and worker characteristics: A French perspective. Working Paper, CREST-INSEE, Paris.
- Crépon, B., and F. Kramarz. 2002. Employed 40 hours or not employed 39: Lessons from the 1982 mandatory reduction of the workweek. *Journal of Political Economy* 110:1355–89.
- Daniel, K., and J. S. Heywood. 2007. The determinants of hiring older workers. *Labor Economics* 14:35–51.
- Danish Minister of Employment. 2003. Agreement among the Liberals, the Conservatives, the Social Democrats, the Socialist People's Party and the Christian Democrats concerning access to the Danish labor market after the enlargement of the European Union on 1 May 2004. Ministry of Employment, Copenhagen.
- Daveri, F., and G. Tabellini. 2000. Unemployment, growth and taxation in industrial countries. *Economic Policy* 15 (30): 47–104.
- Decressin, J., and A. Fátas. 1995. Regional labor market dynamics in Europe. *European Economic Review* 39 (9): 1627–57.

- Devine, T., and N. Kiefer. 1991. *Empirical labor economics: The search approach*. Oxford: Oxford University Press.
- Diamond, P. A. 1971. A model of price adjustment. *Journal of Economic Theory* 3 (2): 150–68.
- Di Nardo J., N. Fortin, and T. Lemieux. 1996. Labor market institutions and the distribution of wages: A semi-parametric approach. *Econometrica* 64:1001–44.
- Disney, R. 2000. The impact of tax and welfare policies on employment and unemployment in OECD countries. IMF Working Paper no. 164, Washington, DC.
- . 2004. Are contributions to public pension programs a tax on employment? *Economic Policy* 19 (39): 267–311.
- Di Tella, R., and R. MacCulloch. 2005. The consequences of labor market flexibility: Panel evidence based on survey data. *European Economic Review* 49 (5): 1225–59.
- Djajic, S., and R. Milbourne. 1988. A general equilibrium model of guestworker migration. *Journal of International Economics* 25:335–51.
- Docquier, F., and H. Rapoport. 2004. Skilled migration: The perspective of developing countries. Policy Research Working Paper Series 3382, World Bank.
- Dolado, J., F. Kramarz, S. Machin, A. Manning, D. Margolis, and C. Teuling. 1996. The economic impact of minimum wages in Europe. *Economic Policy* 23:317–72.
- Dolmas, J., and G. Huffman. 2003. On the political economy of immigration and income redistribution. *International Economic Review* 45:1129–68.
- Dolton, P., and D. O'Neill. 1996. Unemployment duration and the restart effect: Some experimental evidence. *Economic Journal* 106:387–400.
- . 2002. The long-run effects of unemployment monitoring and work-search programs: Experimental evidence from the United Kingdom. *Journal of Labor Economics* 20:381–403.
- Dorn, D., and A. Sousa-Poza. 2005. Early retirement: Free choice or forced decision? Working Paper no. 1542, CESifo, Munich.
- Dostie, B. 2006. Wages, productivity and aging. Working Paper, Institute of Applied Economics, Montreal.
- Dunlop, D. M. 1944. The Karaits of East Asia. *Bulletin of the School of Oriental and African Studies* (University of London) 11 (2): 276–89.
- Dustmann, C., F. Fabbri, and I. Preston. 2005. The impact of immigration on the British labor market. *Economic Journal* 115 (507): 324–42.
- Dustmann, C., and A. Gitz. 2005. *Immigration, jobs and wages: Theory, evidence and opinion*. London: CREAM-CEPR.
- Duval, R. 2003. The retirement effects of old-age pension and early retirement schemes in OECD countries. OECD Economics Department Working Paper no. 370, Paris.
- Dygaló, N. N., and J. M. Abowd. 2005. Estimating experience-productivity profiles from earnings over employment spells. Working Paper, University of Western Ontario.
- Ebbinghaus, B., and J. Visser. 2000. *Trade unions in Western Europe since 1945*. London: Macmillan.
- Eberts, R., and J. A. Stone. 1986. On the contract curve: A test of alternative models of collective bargaining. *Journal of Labor Economics* 4 (1): 66–81.
- Ehrenberg, R. G., and R. S. Smith. 2006. *Modern labor economics, ninth edition*. New York: Addison Wesley.
- Eissa, N. O. 1995. Taxation and labor supply of married women: The tax reform act of 1986 as a natural experiment. NBER Working Paper no. W5023, Washington, DC.

- Eissa, N. O., and J. Liebman. 1996. Labor supply responses to the Earned Income Tax Credit. *Quarterly Journal of Economics* 111 (2): 605–37.
- Elmeskov, J., J. P. Martin, and S. Scarpetta. 1998. Key lessons for labor market reforms: Evidence from OECD countries' experience. *Swedish Economic Policy Review* 5 (2): 205–52.
- Emerson, M. 1998. Regulation or de-regulation of the labour market: Policy regimes for the recruitment and dismissal of employees in industrialised countries. *European Economic Review* (April): 775–817.
- Esping Andersen, G. 1990. *The three worlds of welfare capitalism*. Cambridge: Polity Press.
- Espinosa, M. P., and C. Rhee. 1989. Efficient wage bargaining as a repeated game. *Quarterly Journal of Economics* 104:565–88.
- European Commission. 2005. The economically inactive population in the EU: Out of labor force or potential labor supply? A perspective from the EU Labor Force Survey. In *Employment in Europe 2005*. Directorate General for Economic and Financial Affairs, Brussels.
- Faini, R., J. DeMelo, and K. F. Zimmermann, eds. 1999. *Migration: The controversies and the evidence*. Cambridge: Cambridge University Press.
- Faini, R., and A. Venturini. 1995. Migration and growth: The experience of southern Europe. CEPR Discussion Paper no. 964. London.
- Fassmann, R., and C. Hintermann. 1997. Migrationspotential Ostmitteleuropa: Struktur und Motivation potentieller Migranten aus Polen, der Slowakei, Tschechien und Ungarn. ISR-Scientific Report 15, Innsbruck.
- Flaig, G. 2001. Die Abschätzung der Migrationspotentiale der osteuropäischen EU-Beitrittsländer. *Konjunkturpolitik—Applied Economics Quarterly*, Supplement 52:55–76.
- Flanagan, R. 1999. Macroeconomic performance and collective bargaining: An international perspective. *Journal of Economic Literature* 37 (3): 1150–75.
- Flinn, C. 2002. Interpreting minimum wage effects on wage distributions: A cautionary tale. *Annales d'Économie et de Statistique* 67–68 (July/December): 309–55.
- . 2007. Minimum wage effects on labor market outcomes under search, matching and endogenous contact rates. *Econometrica* 74 (4): 1013–62.
- Frazis, H., and M. A. Loewenstein. 2005. Reexamining the returns to training functional form, magnitude, and interpretation. *Journal of Human Resources* 40:453–76.
- Freeman, R. B., and J. L. Medoff. 1984. *What do unions do?* New York: Basic Books.
- Friedberg, R. M., and J. Hunt. 1995. The impact of immigrants on host countries wages, employment, and growth. *Journal of Economic Perspectives* 9:23–44.
- Friedman, M. 1962. *Capitalism and freedom*. Chicago: University of Chicago Press.
- Fuest, C., and B. Huber. 2000. Is tax progression really good for employment? A model with endogenous hours of work. *Labor Economics* 7:79–93.
- Galasso, V. 2006. *The political future of social security in aging societies*. Cambridge: MIT Press.
- Garibaldi, P., A. Borgarello, and L. Pacelli. 2003. Employment protection legislation and the size of firms: A close look at the Italian case. *Giornale degli Economisti e Annali di Economia*
- Garibaldi, P., J. Konings, and C. Pissarides. 1997. Gross job reallocation and labor market policy. In *Unemployment policy: Government options for the labour market*, ed. D. J. Snower and G. de la Dehesa. Cambridge: Cambridge University Press.
- Garibaldi, P., and P. Mauro. 2002. Anatomy of employment growth. *Economic Policy* 17 (34): 67–114.

- Garibaldi, P., J. Oliveira, and J. C. van Ours. 2008. *Health, longevity and productivity*, forthcoming. Oxford: Oxford University Press.
- Garibaldi, P., and G. Violante. 2005. The employment effects of severance payments with wage rigidities. *Economic Journal* 115 (506): 799–832.
- Gavosto, A., A. Venturini, and C. Villosio. 1999. Do immigrants compete with natives? *Labor* 13 (3): 603–21.
- Geishecker, I., and H. Görg. 2005. The vertical investment controversy: Re-estimating the knowledge-capital model for different types of FDI. Mimeo, University of Nottingham.
- Gielen, A., and J. C. van Ours. 2006. Age-specific cyclical effects in job reallocation and labor mobility. *Labor Economics* 13:493–504.
- Gindling, T. H., and K. Terrell. 2004. Minimum wages and the wages of formal and informal sector workers in Costa Rica. University of Michigan Working Paper 04-102.
- Glyn, A. 2006. *Capitalism unleashed: Finance, globalization and welfare*. Oxford: Oxford University Press.
- Gornick, J. C., and M. K. Meyers. 2003. Support for working families: Work and care policies across welfare states. CESifo DICE Report 4:13–18.
- Gorter, C., and G.R.J. Kalb. 1996. Estimating the effect of counselling and monitoring the unemployed using a job search model. *Journal of Human Resources* 31:590–610.
- Gramlich, E. 1976. Impact of minimum wages on other wages, employment and family incomes. *Brooking Papers on Economic Activity* 2:409–51.
- Graversen, B. K., and J. C. van Ours. 2008. How to help unemployed find jobs quickly: Experimental evidence from a mandatory activation program. *Journal of Public Economics*, forthcoming.
- Gregg, P. 2000. The use of wage floors as policy tools. *OECD Economic Studies* 31: 133–46.
- Gregg, P., and A. Manning. 1997. Skill-biased change, unemployment and inequality. *European Economic Review* 41:1173–1200.
- Gregg, P., and J. Waldfogel. 2005. Symposium on parental leave, early maternal employment and child outcomes: Introduction. *Economic Journal* 115:F1–F6.
- Grout, P. 1984. Investment and wages in the absence of binding contracts: A Nash bargaining approach. *Econometrica* 52:449–60.
- Grubb, D. 2000. Eligibility criteria for unemployment benefits. *OECD Employment Outlook* 31: 147–84.
- Grubb, D., and W. Wells. 1997. Employment regulations and patterns of work in EC countries. *OECD Economic Studies* 21: 7–58.
- Gruber, J. 1997. The incidence of payroll taxation: Evidence from Chile. *Journal of Labor Economics* 15 (3): S72–S101.
- Gruber, J., and D. A. Wise. 1997. Social security programs and retirement around the world. NBER Working Paper no. 6134, Cambridge, MA.
- Hamermesh, D. S. 1980. Social insurance and consumption: An empirical inquiry. *American Economic Review* 72 (1): 101–13.
- . 1995. Myth and measurement: The new economics of the minimum wage; Comment. *Industrial and Labor Relations Review* 20 (4): 830–34.
- . 1996. *Labor demand*. Princeton, NJ: Princeton University Press.
- Hansen, C. 1999. Lower tax progression, longer hours and higher wages. *Scandinavian Journal of Economics* 101:49–65.
- Hansen, J. D. 2003. Immigration and income redistribution in welfare states. *European Journal of Political Economy* 19:735–46.

- Hanson G. H., R. Robertson, and A. Spilimbergo. 2002. Does border enforcement protect U.S. workers from illegal immigration? *Review of Economics and Statistics* 84 (1): 73–92.
- Hanson, G. H., and M. J. Slaughter. 1999. The Rybczynski theorem, factor price equalization, and immigration: Evidence from U.S. states. NBER Working Paper no. 7074, Cambridge, MA.
- Hanushek, E. A. 2002. Publicly provided education. In *Handbook of Public Economics*, ed. A. Auerbach and M. Feldstein, 4:2045–41. Amsterdam: Elsevier.
- Harris, J. R., and M. P. Todaro. 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review* 60:126–42.
- Hatton, T. J. 1995. A model of U.K. emigration, 1870–1913. *Review of Economics and Statistics* 77:407–15.
- . 2004. Seeking asylum in Europe. *Economic Policy* 19:5–62.
- Hatton, T. J., and J. Williamson. 1998. *The age of mass migration: Causes and economic impact*. New York: Oxford University Press.
- Heckman, J. J., R. J. LaLonde, and J. A. Smith. 1999. The economics and econometrics of active labor market programs. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 1865–2097. Amsterdam: Elsevier.
- Heijdra, B., C. Keuschnigg, and W. Kohler. 2002. Eastern enlargement of the EU: Jobs, investment and welfare in the present EU member countries. CESifo Working Paper no. 718, Munich.
- Hellerstein, J. K., N. David, and K. R. Troske. 1999. Wages, productivity and worker characteristics: Evidence from plant-level production function and wage equations. *Journal of Labor Economics* 17:409–46.
- Helpman, E. 1984. A simple theory of trade with multinational corporations. *Journal of Political Economy* 92:451–71.
- Hicks, J. R. 1932. *The theory of wages*. London: Macmillan.
- Hille, H., and T. Straubhaar. 2001. The impact of EU-enlargement on migration movements and economic integration: Results of recent studies. In *Migration policies and EU-enlargement*, ed. OECD, 79–100. Paris: OECD.
- Hirschman, A. O. 1970. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Cambridge, MA: Harvard University Press.
- Holden, S., and O. Raaum. 1991. Wage moderation and union structure. *Oxford Economic Papers* 43:409–23.
- Holmlund, B. 1998. Unemployment insurance in theory and practise. *Scandinavian Journal of Economics* 100 (1): 113–41.
- Holzer, H. 1986. Reservation wages and their labor market effects for black and white male youth. *Journal of Human Resources* 21: 157–77.
- Home Office. 2004a. Consequences of EU enlargement. Memorandum submitted to the House of Commons, United Kingdom Parliament, London.
- . 2004b. New figures show accession workers working for the UK. Home Office Press Releases, Reference: 351/2004, London, December 10.
- Hopenhayn, H. A., and J.-P. Nicolini. 1997. Optimal unemployment insurance. *Journal of Political Economy* 105 (2): 412–38.
- Horstmann, I. J., and J. R. Markusen. 1987. Strategic investments and the development of multinationals. *International Economic Review* 28:109–21.
- Hosios, A. J. 1990. On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies* 57:279–98.

- Hunt, J. 1992. The impact of the 1962 repatriates from Algeria on the French labor market. *Industrial and Labor Relations Review* 45 (April): 556–72.
- . 1999. Has work-sharing worked in Germany? *Quarterly Journal of Economics* 114:117–48.
- . 2006. Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association* 4 (5):1014–37.
- Hunya, G. 2004. Employment effects of FDI in new EU members. Mimeo, WIIW, Vienna.
- Hutchens, R. 1986. Delayed payment contracts and a firm's propensity to hire older workers. *Journal of Labor Economics* 4:439–57.
- Ilmakunnas, P., and M. Maliranta. 2005. Technology, labor characteristics and wage-productivity gaps. *Oxford Bulletin of Economics and Statistics* 67:623–45.
- International Organization for Migration. 1998. *Migration potential in central and Eastern Europe*. Geneva: Technical Cooperation Centre for Europe and Central Asia, International Organization for Migration (IOM).
- Isacson, I. 1999. Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics* 6:471–89.
- Jackman, R., R. Layard, and S. Nickell. 1996. Combatting unemployment: Is flexibility enough? CEP Discussion Paper no. 0293.
- Jaumotte, F. 2003. Female labor force participation: Past trends and main determinants in OECD countries. OECD Economics Department Working Paper no. 376, Paris.
- Johnson, P. 1993. Ageing and European economic demography. In *Labour markets in an ageing Europe*, ed. P. Johnson and K. F. Zimmermann. Cambridge: Cambridge University Press.
- Jones, P. 1997. The impact of the minimum wage legislation in developing countries where coverage is incomplete. WP/98-2, Institute of Economics and Statistics, Oxford.
- Jones, S.R.G., and C. J. McKenna. 1994. A dynamic model of union membership and employment. *Economica* 61:179–89.
- Jones, S.R.G., and W. C. Riddell. 1999. The measurement of unemployment: An empirical approach. *Econometrica* 67:147–62.
- Judson, R. A., and A. K. Owen. 1999. Estimating dynamic panel data models: A guide for macroeconomists. *Economic Letters* 65:9–15.
- Kahn, L. M. 1998. Collective bargaining and the interindustry wage structure: International evidence. *Economica* 65 (260): 507–34.
- . 2000. Wage inequality, collective bargaining, and relative employment from 1985 to 1994: Evidence from fifteen OECD countries. *Review of Economics and Statistics* 82 (4): 564–79.
- Kaitz, H. 1970. Experience of the past: The national minimum, in youth unemployment and minimum wages. Bulletin no. 1657, 30–54, U.S. Department of Labor, Bureau of Labor Statistics, Washington, DC.
- Kapteyn, A., A. Kalwij, and A. Zaidi. 2004. The myth of work-sharing. *Labour Economics* 11:293–313.
- Katz, L. F., and B. D. Meyer. 1990. Unemployment insurance, recall expectations, and unemployment outcomes. *Quarterly Journal of Economics* 105 (4): 973–1002.
- Kaufman, B., and J. Hotchkiss. 2006. *The economics of the labor market*. 7th ed. Thomson South-Western.
- Kennan, J. 1995. The elusive effects of minimum wages. *Journal of Economic Literature* 33 (4): 1950–65.

- Keuschnigg, C., and W. Kohler. 2002. Eastern enlargement of the EU: How much is it worth for Austria? *Review of International Economics* 10:324–42.
- Kiefer, N. 1988. Economic duration data and Hatard Functions. *Journal of Economic Literature* 26:646–79.
- Kiviet, J. 1995. On bias, inconsistency and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68:53–78.
- Klepinger, D. H., T. R. Johnson, and J. M. Joesch. 2002. Effects of unemployment insurance work-search requirements: The Maryland experiment. *Industrial and Labor Relations Review* 56:3–22.
- Kluve, J. 2006. The effectiveness of European active labor market policy. Working Paper no. 2018, IZA, Bonn.
- Kluve, J., and C. M. Schmidt. 2002. Can training and employment subsidies combat European unemployment? *Economic Policy* 35:411–48.
- Kohler, W. 2003. Eastern enlargement of the EU: A comprehensive welfare assessment. Working Paper no. 0316, Department of Economics, Johannes Kepler University, Linz.
- Kramarz, F., P. Cahuc, B. Crépon, T. Schank, O. Nordstrom Skans, and G. van Lomwel. 2008. Work sharing. In 2007. *Working Hours and Job Sharing in the EU and USA*, ed. T. Boeri, M. C. Burda, and F. Kramarz. Oxford: Oxford University Press.
- Krieger, T. 2003. Migration trends in an enlarged EU. European Foundation for the Improvement of Working and Living Conditions, Dublin.
- Krueger, A., and B. Meyer. 2002. Labor supply effects of social insurance. In *Handbook of public economics*, ed. A. Auerbach and M. Feldstein. Amsterdam: Elsevier.
- Krugman, P. 1991. Increasing returns and economic geography. *Journal of Political Economy* 99:483–99.
- Kugler, A., and G. Saint-Paul. 2000. Hiring and firing costs, adverse selection and long-term unemployment, Working Paper no. 447, Universitat Pompeu Fabra.
- Kvist, J. 2004. Does EU enlargement start a race to the bottom? Strategic interaction among EU member states in social policy. *Journal of European Social Policy* 14 (3): 301–18.
- Lalive, R., J. C. van Ours, and J. Zweimüller. 2005. The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association* 3 (6): 1386–1417.
- . 2006. How changes in financial incentives affect the duration of unemployment. *Review of Economic Studies* 73:1009–38.
- Lalive, R., and J. Zweimüller. 2004. Benefit entitlement and unemployment duration: The role of policy endogeneity. *Journal of Public Economics* 88 (12): 2587–616.
- Lancaster, T. 1979. Econometric methods for the duration of unemployment. *Econometrica* 47 (4): 939–56.
- Layard, R., O. Blanchard, R. Dornbusch, and P. Krugman. 1992. *East-west migration: The alternatives*. Cambridge, MA: MIT Press.
- Layard, R., S. Nickell, and R. Jackman. 2005. *Unemployment*. New York: Oxford University Press.
- Lazear, E. P. 1979. Why is there mandatory retirement? *Journal of Political Economy* 87:1261–84.
- . 1990. Job security provisions and unemployment. *Quarterly Journal of Economics* 3:699–726.
- Leamer, E., and J. Levinsohn. 1995. International trade theory: The evidence. In *Handbook of International Economics*, ed. G. Crossman and K. Rogoff, 3:1339–94. Amsterdam: Elsevier.

- Lechner, M., and J. Smith. 2007. What is the value added by caseworkers? *Labour Economics* 14:135–51.
- Lerner, A. P. 1952. Factor prices and international trade. *Economica* 19:1–15.
- Lester, R. 1947. Marginalism, minimum wages, and labor markets. *American Economic Review* 37 (1): 135–48.
- Leuven, E. 2005. The economics of private sector training: A survey of the literature. *Journal of Economic Surveys* 19:91–111.
- Levine, P. 1999. The welfare economics of migration control. *Population Economics* 12: 23–43.
- Lewis, E. 2005. Immigration, skill mix, and the choice of technique. Working Paper no. 05-8, Federal Reserve Bank of Philadelphia.
- Lippman, S. A., and J. J. McCall, eds. 1979. *Studies in the economics of search*. Amsterdam: North-Holland.
- Ljungqvist, L., and T. J. Sargent. 2003. European unemployment and turbulence revisited in a matching model. <http://homepages.nyu.edu/~ts43/RESEARCH/researchframe.htm>.
- Loewenstein, M. A., and J. R. Spletzer. 1998. Dividing the costs and returns to general training. *Journal of Labor Economics* 16 (1): 142–71.
- Löffelholz, H.-D., and G. Köpp. 1998. *Ökonomische Auswirkungen der Zuwanderung nach Deutschland*. Schriftenreihe des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung, New Series, no. 63. Berlin: Duncker und Humblot.
- Ludema, R., and I. Wooton. 1999. Regional integration, trade and migration: Are demand linkages relevant in Europe? In *Migration: The controversies and the evidence*, ed. R. Faini, J. DeMelo, and K. F. Zimmermann. Cambridge: Cambridge University Press.
- Lynch, L. 1983. Job search and youth unemployment. *Oxford Economic Papers* 35: 271–82.
- Maddison, A. 1995. *Monitoring the world economy, 1820–1992*. Paris: OECD.
- Maloney, W. F., and J. N. Mendez. 2003. Measuring the impact of minimum wages: Evidence from Latin America. NBER Working Papers 9800, National Bureau of Economic Research, Cambridge, MA.
- . 2004. Minimum wages in Latin America. In *Law and employment: Lessons from Latin American and the Caribbean*, ed. J. Heckman and C. Pages. Chicago: University of Chicago Press.
- Manning, A. 2003. *Monopsony in motion*. Princeton, NJ: Princeton University Press.
- Manning, A., and B. Petrongolo. 2006. The part-time pay penalty. CEP Discussion Papers 0679, Centre for Economic Performance, London School of Economics.
- Marimon, R., and F. Zilibotti. 1999. Unemployment vs. mismatch of talents: Reconsidering unemployment benefits. *Economic Journal* 109 (455): 266–91.
- . 2000. Employment and distributional effects of restricting working time. *European Economic Review* 44 (7): 1291–1326.
- Markusen, J. R. 1983. Factor movements and commodity trade as complements. *Journal of International Economics* 13:341–56.
- . 1984. Multinationals, multi-plant economies and the gains from trade. *Journal of International Economics* 16:205–26.
- . 2002. *Multinational firms and the theory of international trade*. Cambridge, MA: MIT Press.
- Markusen, J. R., and A. J. Venables. 1998. Multinational firms and the new trade theory. *Journal of International Economics* 46:183–203.

- Martin, J. P., and D. Grubb. 2001. What works and for whom: A review of OECD countries' experience with active labor market policies. Working Paper no. 2001:14, OECD, Paris.
- Mayda, A., and D. Rodrik. 2001. Why are some people (and countries) more protectionist than others? NBER Working Paper no. 8461, Cambridge, MA.
- McConnell, C. R., S. L. Brue, and D. Macpherson. 2008. *Contemporary labor economics*, eighth edition. Boston: McGraw-Hill Irwin.
- McCurdy, T. E., and J. Pencavel. 1986. Testing between competing models of wage and employment determination in unionized markets. *Journal of Political Economy* 94: S3–S9.
- McDonald, I. M., and R. M. Solow. 1981. Wage bargaining and employment. *American Economic Review* 71:896–908.
- McLennan, K. 2005. What do unions do? A management perspective. *Journal of Labor Research* 26(4):641–88.
- Metzler, A., and S. Richard. 1981. A rational theory of the size of government. *Journal of Political Economy* 89 (51): 914–27.
- Meyer, B. D. 1990. Unemployment insurance and unemployment spells. *Econometrica* 58(4):757–82.
- . 1995a. Natural and quasi-experiments in science. *Journal of Business and Economic Statistics* 13(2):151–61.
- . 1995b. Lessons from the U.S. unemployment experiments. *Journal of Economic Literature* 33(1):91–131.
- Miles, T. 2000. Common law exceptions to employment at will and US labor markets. *Journal of Law, Economics and Organization* 16 (1): 74–101.
- Mincer, J. 1976. Unemployment effects of minimum wages. *Journal of Political Economy* 84:87–104.
- Moene, E. 1997. Competitive search equilibrium. *Journal of Political Economy* 105:385–411.
- Mortensen, D. T. 1977a. Specific Human Capital and Labor Turnover. Discussion Papers 289, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- . 1977b. Unemployment insurance and labor supply decisions. Discussion Papers 271, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- . 2003. *Wage dispersion: Why are similar workers paid differently?* Cambridge, MA: MIT Press.
- Mundell, R. 1957. International trade and factor mobility. *American Economic Review* 47:321–35.
- Narendranathan, W., S. Nickell, and J. Stern. 1985. Unemployment benefit—revisited. *Economic Journal* 95 (378): 307–29.
- Nash, J. 1950. The bargaining problem. *Econometrica* 18(2):155–62.
- . 1953. Two-person cooperative games. *Econometrica* 21(1):128–40.
- Naylor, R., and M. Cripps. 1993. An economic theory of the open shop. *European Economic Review* 37:1599–1620.
- Naz, G. 2004. The impact of cash-benefit reform on parents' labor force participation. *Journal of Population Economics* 17:369–83.
- Neumark, D., M. Schweitzer, and W. Wascher. 2004. Minimum wage effects throughout the wage distribution. *Journal of Human Resources* 39 (2): 425–50.

- Neumark, D., and W. Stock. 1999. Age discrimination law and labor market efficiency. *Journal of Political Economy* 107:1081–1125.
- Neumark, D., and W. Wascher. 2000. Minimum wages: A case study of the fast food industry in New Jersey and Pennsylvania; Comment. *American Economic Review* 90: 1362–96.
- . 2007. Minimum wages and employment. IZA Discussion Paper no. 2570, Bonn.
- Nickell, S. 1979. The effects of unemployment and related benefits on the duration of unemployment. *Economic Journal* 89 (353): 34–49.
- . 1981. Biases in dynamic models with fixed effects. *Econometrica* 49:1417–26.
- . 1997. Unemployment and labor market rigidities: Europe versus North America. *Journal of Economic Perspectives* 11 (3): 55–74.
- . 2006. Work and taxes. In *Tax policy and labor market Performance*, ed. J. Agell and P. B. Sorensen. Cambridge, MA: MIT Press.
- Nickell, S., and M. Andrews. 1983. Unions, real wages and employment in Britain, 1951–79. *Oxford Economic Papers* 35 (1): 183–206.
- Nickell, S., and R. Layard. 1999. Labor market institutions and economic performance. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 3B: 3029–84. Amsterdam: Elsevier.
- Nickell, S., L. Nunziata, and W. Ochel. 2005. Unemployment in the OECD since the 1960s: What do we know? *Economic Journal* 115: 1–27.
- Nickell, S., and J. van Ours. 2000. The Netherlands and United Kingdom: A European unemployment miracle? *Economic Policy* 15 (30): 135–80.
- Oaxaca, R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* (Department of Economics, University of Pennsylvania, and Osaka University Institute of Social and Economic Research Association) 14 (3): 693–709.
- OECD. 1994. *Jobs study*. Paris: OECD.
- . 1995. *Employment outlook*. Paris: OECD.
- . 1998. Making the most of the minimum: Statutory minimum wages, employment, and poverty. In *Employment outlook*, 31–77. Paris: OECD.
- . 1999. *Employment outlook*. Paris: OECD.
- . 2001a. *Employment outlook*. Paris: OECD.
- . 2001b. *International migration of the highly skilled*. Paris: OECD.
- . 2002. *Employment outlook*. Paris: OECD.
- . 2003. *Employment outlook*. Paris: OECD.
- . 2004a. *Benefits and wages*. Paris: OECD.
- . 2004b. *Employment outlook*. Paris: OECD.
- . 2005a. *Education at a glance*. Paris: OECD.
- . 2005b. *Employment outlook*. Paris: OECD.
- . 2006a. *Boosting jobs and incomes: Policy lessons from reassessing the OECD jobs strategy*. Paris: OECD.
- . 2006b. *Employment outlook*. Paris: OECD.
- . 2006c. *Jobs study*. Paris: OECD.
- . 2006d. *Live longer, work longer*. Paris: OECD.
- . 2007. *Pensions at a glance*. Paris: OECD.
- Osborne, M., and A. Rubinstein. 1990. *Bargaining and markets*. San Diego: Academic Press.
- Pavoni, N., and G. Violante. 2007. Optimal welfare-to-work programs. *Review of Economic Studies* 74 (1): 283–318.

- Pencavel, J. 2003. The surprising retreat of union Britain. NBER Working Papers no. 9564, National Bureau of Economic Research, Cambridge, MA.
- Persson, T., and G. Tabellini. 2000. *Political economics: Explaining economic policy*. Cambridge, MA: MIT Press.
- Pestieau, P. 2003. Ageing, retirement and pension reforms. *World Economy* 26:1447–57.
- Piketty, T. 1998. L'impact des incitations financières au travail sur les comportements individuels: Une estimation pour le cas français. *Economie et Prévision* 132–133:1–36.
- Pissarides, C. 1979. Job matchings with state employment agencies and random search. *Economic Journal* 89:818–33.
- . 1998. The impact of employment tax cuts on unemployment and wages: The role of unemployment benefits and tax structure. *European Economic Review* 42:155–83.
- Portugal, P., and A. R. Cardoso. 2001. Disentangling the minimum wage puzzle: An analysis of job accession and separation from a longitudinal matched employer employee data set. CEPR Discussion Paper no. 2844, London.
- Postel-Vinay, F., and A. Clark. 2006. Job security and job protection. IZA Discussion Paper no. 1489, Bonn.
- Puhanyi, P. A. 2001. Is labor mobility an adjustment mechanism for Euroland? Empirical evidence for western Germany, France, and Italy. *German Economic Review* 2:127–40.
- Rangazas, P. 2002. The quantity and quality of schooling and U.S. labor productivity growth (1870–2000). *Review of Economic Dynamics* 5:932–64.
- Ransom, R. L., and R. Sutch. 1986. The labor of older Americans: Retirement of men on and off the job, 1870–1937. *Journal of Economic History* 46:1–30.
- Razin, A., E. Sadka, and P. Swagel. 2002. Tax burden and migration: A political economy theory and evidence. *Journal of Public Economics* 85 (2): 167–90.
- Rebitzer, J. B., and L. J. Taylor. 1995a. The consequences of minimum wage laws: Some new theoretical ideas. *Journal of Public Economics* 56 (2): 245–55.
- . 1995b. Do labor markets provide enough short-hour jobs? An analysis of work hours and work incentives. *Economic Inquiry* 33 (2): 257–73.
- . 1995c. Efficiency wages and employment rents: The employer-size wage effect in the job market for lawyers. *Journal of Labor Economics* 13 (4): 678–708.
- Remery, C., K. Henkens, J. Schippers, and P. Ekamper. 2003. Managing an aging workforce and a tight labor market: Views held by Dutch employers. *Population Research and Policy Review* 22:21–40.
- Robinson, C. 1989. The joint determination of union status and union wage effects: Some tests of alternative models. *Journal of Political Economy* 97 (3): 639–67.
- Rodrik, D. 1998. Why do more open economies have bigger governments? *Journal of Political Economy* 106 (5): 997–1032.
- . 1999. Final remarks. In *Migration: The controversies and the evidence*, ed. R. Faini, J. DeMelo, and K. F. Zimmermann, 314–17. Cambridge: Cambridge University Press.
- Rogerson, R. 1997. Theory ahead of language in the economics of unemployment. *Journal of Economic Perspectives* 11 (1): 73–92.
- Roodenburg, H., R. Euwals, and H. ter Rele. 2003. *Immigration and the Dutch economy*. The Hague: CPB and Koninklijke De Swart.
- Rosen, S. 1974. Hedonic prices and implicit markets. *Journal of Political Economy* 82: 34–55.
- Roy, A. D. 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3:314–17.

- Ruhm, C. J. 1998. The economic consequences of parental leave mandates: Lessons from Europe. *Quarterly Journal of Economics* 285–317.
- Saez, E. 2002. Optimal income transfer programs: Intensive versus extensive labor supply responses. *Quarterly Journal of Economics* 1039–73.
- Saint-Paul, G. 1993. On the political economy of labor market flexibility. CEPR Discussion Papers 803. Other version: On the political economy of labor market flexibility. DELTA Working Papers 93-02, DELTA (Ecole normale supérieure).
- . 1997. *Dual labor market: A macroeconomic perspective*. Cambridge, MA: MIT Press.
- . 2000. *The political economy of labor market institutions*. Oxford: Oxford University Press.
- Salop, S. 1979. A model of the natural rate of unemployment. *American Economic Review* 74:117–25.
- Samuelson, P. A. 1949. International factor-price equalization once again. *Economic Journal* 59:181–97.
- Scarpetta, S. 1996. Assessing the role of labor-market policies and institutional factors on unemployment: A cross-country study. *OECD Economic Studies* 26:43–98.
- . 1999. Labor market reforms and unemployment: Lessons from the experience of the OECD countries. Inter-American Development Bank Working Papers 382, Inter-American Development Bank, Washington, DC.
- Scheve, K., and M. Slaughter. 2001. Labor-market competition and individual preferences over immigration policy. *Review of Economics and Statistics* 83 (1): 133–45.
- Schneider, F., and D. Enste. 2000. Shadow economies: Size, causes, and consequences. *Journal of Economic Literature* 38 (1): 77–114.
- Shapiro, C., and J. Stiglitz. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74:433–44.
- Sinn, H.-W. 2004. Migration and social replacement incomes: How to protect low-income workers in the industrialized countries against the forces of globalization and market integration. Plenary lecture given at the IIPF Congress, Milan, August.
- Sjaastad, L. A. 1962. The costs and returns of human migration. *Journal of Political Economy* 70 (5): 80–93.
- Smith, J. P., and B. Edmonston. 1997. *The new Americans: Demographic and fiscal effects of immigration*. Washington, DC: National Academy Press.
- Snowder, D., and G. de la Dehesa, eds. 1996. *Unemployment policy*. Cambridge: CEPR and Cambridge University Press.
- Sobel, R. S. 1999. Theory and evidence on the political economy of the minimum wage. *Journal of Political Economy* 107 (4): 761–85.
- Sociaal en Cultureel Planbureau. 1998. *Sociaal and Cultureel Rapport, 1998*. The Hague: Staatsuitgeverij.
- Soskice, D., and T. Iversen. 2000. The non-neutrality of monetary policy with large price or wage setters. *Quarterly Journal of Economics* 115 (1): 265–84.
- Stewart, M. 2004. The employment effects of the national minimum wage. *Economic Journal* 114:C110–C116.
- Stigler, G. 1946. The economics of minimum wage legislation. *American Economic Review* 36:535–43.
- Stock, J. H., and D. A. Wise. 1990. Pensions: The option value of work and retirement. *Econometrica* 58:1151–80.

- Tanaka, S. 2005. Parental leave and child health across OECD countries. *Economic Journal* 115:F7–F28.
- Topel, R. H. 1986. Local labor markets. *The Journal of Political Economy* 54:111–43.
- Trabold, H., and P. Trübswetter. 2001. Schätzung der Beschäftigungs- und Lohneffekte der Zuwanderung. In *Migration: Potential und Effekte für den deutschen Arbeitsmarkt*, Final Report for the Hans-Böckler-Foundation, ed. M. Brücker, H. Trabold, P. Trübswetter, and C. Weise. Berlin: German Institute for Economic Research (DIW).
- U.S. Immigration and Naturalization Service. 1996. *Statistical Yearbook of the Immigration and Naturalization Service*. Washington, DC.
- van den Berg, G. J. 1990a. Nonstationarity in job search theory. *Review of Economic Studies* 57:255–77.
- . 1990b. Search behavior, transitions to nonparticipation and the duration of unemployment. *Economic Journal* 100:842–65.
- van den Berg, G. J., and B. van der Klaauw. 2006. Counseling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment. *International Economic Review* 47:895–936.
- van den Berg, G. J., B. van der Klaauw, and J. C. van Ours. 2004. Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics* 22:211–41.
- van Lomwel, A.G.C., and J. C. van Ours. 2005. On the employment effects of part-time labor. *De Economist* 153 (4): 451–60.
- van Ours, J. C. 2004a. The locking-in effect of subsidized jobs. *Journal of Comparative Economics* 32 (1): 37–52.
- . 2004b. A pint a day raises a man's pay; but smoking blows that gain away. *Journal of Health Economics* 23 (5): 863–86.
- van Ours, J. C., and M. Vodopivec. 2006. How shortening the potential duration of unemployment benefit's affects the duration of unemployment: Evidence from a natural experiment. *Journal of Labor Economics* 24 (2): 351–78.
- . 2008. Does reducing unemployment insurance generosity reduce job match quality? *Journal of Public Economics*, forthcoming.
- Venables, A. 1999. Trade liberalisation and factor mobility: An overview. In *Migration: The controversies and the evidence*, ed. R. Faini, J. DeMelo, and K. F. Zimmermann, 23–47. Cambridge: Cambridge University Press.
- Vodopivec, M. 2004. *Income support for the unemployed*. Washington, DC: World Bank.
- von Wachter, T. 2002. The end of mandatory retirement in the US: Effects on retirement and implicit contracts. Center for Labor Economics Working Paper no. 49, University of California, Berkeley.
- Walker, A. 1998. Adjusting to an aging workforce in Europe—Policy and practice. Aging of the Workforce workshop, Brussels, March 23–24.
- Warr, P. 1998. Aging, competence and learning at work. Aging of the Workforce workshop, Brussels, March 23–24.
- Wasmer, E. 2006. The economics of Prozac: Do employees really gain from strong employment protection? IZA DP no. 2460, Bonn.
- Welch, F. 1976. Minimum wage legislation in the United States. In *Evaluating the labor market effects of social programs*, ed. O. Ashenfelter and J. Blum. Princeton, NJ: Princeton University Press.
- Williamson, O. E. 1975. *Markets and hierarchies, analysis and antitrust implications: A study in the economics of internal organization*. New York: Free Press.

- Winter-Ebmer, R., and J. Zweimüller. 1996. Immigration and the earnings of young native workers. *Oxford Economic Papers* 48:473–91.
- Wright, R. 1986. The redistributive roles of unemployment insurance and the dynamics of voting. *Journal of Public Economics* 31:377–99.
- Zavodny, M. 2000. The effect of the minimum wage on employment and hours. *Labor Economics* 7:729–50.
- Zimmermann, K. F., ed. 2005. *European migration—What do we know?* Oxford: Oxford University Press.

This page intentionally left blank

Index

- Activation: of active labor market policies, 255–56; of unemployed workers, 273–76
- Active labor market policies, 255–76, 277; activation and, 255–56; empirical evidence on, 263; experimental studies on, 263–68; measures and cross-country comparisons for, 256–58; need for, 271–72; nonexperimental studies on, 268–70; policy issues in, 270–71; for public employment services, 255, 271; reason for existence of, 272; subsidized employment and, 255; theory of, 258–62; training and, 255; for unemployed workers, 273–76, 282–83
- Admission procedures, migration policies and, 175
- Advance notice period in employment protection legislation, 199
- Adverse selection: family policies and, 153; labor market institutions and, 20; unemployment benefits and, 244
- Aggregate labor supply, 7–9; stated reservation wage and, 8–9
- Allocation Parentale d'Éducation (APE), 146–47
- Annual hours of work, 110
- Asylum seekers, migration policies and, 175–76
- Average tax rates, 82
- Average wages, 82–86; ratio of minimum to, 30–31
- Bargaining coordination: unemployment and, 71; union density and, 71
- Bargaining models, 56–57
- Beveridge curve, 222, 258–62
- Bilateral monopoly, 36
- Brain-drain effect of migration, 194
- Centralization of collective bargaining, 56
- Child care: arrangements for, 139; costs for provision of, 142; facilities for, 142, 144, 154–55; hours of work and, 154–55; responsibilities for, 146; subsidies for, 153
- Citizenship, migration policies and, 175
- Collective bargaining, 3, 15, 20; agreements on, 277–78; reduction of working hours in, 110; regulation of working hours in, 116–17; centralization of, 56; minimum wages and, 36; reasons for decentralizing, 74–75; right-to-manage structure of, 59–60; unions and, 51–79; working hours regulation and, 101, 116–17
- Compensating wage differentials, 63–65
- Competitive labor market, 4, 24–25. *See also* Labor market; equilibrium in, 11; labor supply curve in, 88; migration policies and, 179–80; minimum wage in, 33; payroll taxes in, 86, 87, 98–100; unemployment benefits and, 230–34
- Compulsory schooling age, 15, 171
- Contract curve, 60
- Contracts: efficient, 60; employment protection legislation and, 204–5
- Cross-border worker flows, restrictions on, 175
- Defined benefit pension, 122, 124
- Defined contribution pension, 122, 125
- Deunionization, 53
- Difference-in-differences, xv, 278; effect of migration on wages and employment of natives, 187–88; employment protection legislation and, 214; family policies and,

- Difference-in-differences (*Continued*)
 147; reduction in work hours and, 113; in reform analysis, 241; in within-countries studies, 213
- Difference-in-differences estimators, 40
- Discouraged workers, 12
- Displaced workers, return to new job, 129
- Displacement effects, 263
- Dual labor markets, minimum wages in, 37, 45
- Dual-track reform strategy, 202–3
- Duration analysis, 240
- Early childhood education and care, 139
- Early retirement programs, 121–38; employment and, 125–32; mandatory retirement age and, 134–35; measures and cross-country comparisons in, 122–24; optimal retirement age and, 137–38; productivity and, 132–34; reasons for existence of, 136; reasons for phasing out, 135–36; theory on, 124–25
- Earned Income Tax Credit (EITC), 95
- Education, 157–74. *See also* Schooling; Training; government subsidies for in-company, 170; market failures concerning, 158; measures and cross-country comparisons, 158–60; on-the-job, 169–70; productivity and, 157, 163, 171, 174; public spending on, 157; reasons for government provided, 171; retirement programs and, 287–88
- Efficiency wages, unemployment benefits and, 236–37; employment protection and, 217
- Efficient bargaining: unions and, 60, 61, 75, 77
- Employment protection legislation (EPL), 16–17, 20, 158, 199–224, 277; administrative procedures and, 199; advance notice period and, 199; behavior of workers and, 222; contract undoing, 205–6; cross-country analyses of, 211–13; defined, 199; difference-in-differences approach and, 214; efficiency wages and, 217–18; endogeneity of, 214–15; matching, 222; measures and cross-country comparisons in, 200–204; minimum wages and, 277, 278; neutrality and, 204–6; presence of strict, 62; procedural costs and, 199; productivity of workers and, 217; reasons for existence of, 220; reforms in, 21; regulation of working hours by, 101; removing risk neutrality in, 206; retirement programs and, 122, 124, 130, 135; rigid wages and, 206–9; severance payments and, 199; taxes and, 199, 209–10; transfer component and, 199; trial costs and, 199; two-tier regimes in, 210–11; unemployment benefits and, 280–82; unions and, 74, 284–85; wage bargaining in, 223; working hours regulation and, 101
- Entitlement effect, unemployment benefits and, 237–38
- European Trade Unions Confederation (ETUC), 51
- Excess coverage of trade unions, 52, 54
- Experience rating, 249, 281
- Family allowances, 15
- Family day care, 153
- Family income, fertility and, 149
- Family policies, 139–55; balance between work and family life and, 150–51; child care facilities in, 142, 154–55; empirical evidence in, 144–50; hours of work and, 154–55; institutions in, 81; measures and cross-country comparisons in, 140–41; parental leave in, 142–44; reasons for existence of, 153; regulation of working hours and, 101; trade-off between fertility and employment, 151–53
- Fertility: factors affecting, 149; total, 149*n*; trade-off between employment and, 151–53
- Firm-size firm-wage effect, 64–65
- Fiscal substitution effects, 263
- Free-rider problem of unions, 54
- Globalization, labor demand and, 21–28, 289
- Guaranteed minimum income scheme, 183
- Holdup problem, 73–74
- Hours of work. *See* Work hour regulation
- Hours restrictions with and without reservation wage, 6–7

- Household. *See* Family policies
- Human capital, 157; theory of, 169. *See also* Education
- Identical twins, returns to schooling and, 168
- Immigration policies. *See* Migration policies
- In-company training: government subsidies for, 170; provision of, 158
- Job search, 234; formal, 266; informal, 266; theory of, 234–36; unemployment benefits and, 234–36, 249–50
- Kaitz index*, 31*n*
- Labor demand: effect of payroll taxes on, 86; globalization and, 289; inverse, 10; part-time work and, 107–9; wage determination and, 9–10; wedge between labor supply and, 26–27
- Labor market slack: porous participation borders and, 12–13; unemployment benefits and participation in, 251–52
- Labor market institutions, 3, 14, 25–26, 277–90; economic performance and, 1; education and retirement programs and, 287–88; functions performed by, 20–21; future of, 288–90; interactions in, 17; lack of coordination among, 277–78; from normative standpoint, 278; from positive standpoint, 278; prices and, 14–15; quantities and, 15–17; reasons for existence of, 18–21; reforms of, 22; taxes and, 278–80, 285–86; unemployment benefits and, 280–83, 285–86; unions and, 278–80, 284–85
- Labor supply: aggregate, 7–9; in competitive market, 88; effective, 15; effect of payroll taxes on, 86; neoclassical theory of, 235; part-time work and, 107–9; reservation wages and, 4–9; wedge between labor demand and, 26–27
- Labor taxes. *See* Payroll taxes
- Laffer curve, 92*n*
- Late-career job loss, economic consequences of, 129
- Length of stay, migration policies and, 175
- Lighthouse effects, 39
- Locking-in effect, 269
- Low-productivity workers, unemployment among, 277
- Low-wage-cost curve, 104–6
- Lump-sum taxes and transfers, 19
- Mandatory retirement age, 121; elimination of, 131–32; reasons for increasing, 134–35
- Mandatory schooling age. *See* Compulsory schooling age
- Marginal effective tax rate, 82
- Marginal hiring cost of monopsonist, 34–35
- Marginal tax rate, 82
- Maternity leave. *See* Parental leave
- Means testing: unemployment benefits and, 226; welfare benefits in, 92–93
- Migration policies, 175–97; admission procedures and, 175; asylum seekers and, 175–76; brain-drain effect of, 194; citizenship and, 175; competitive labor market and, 179–80; cross-country comparisons in, 176–78; economy with wage rigidities and, 180–81; effects on income distribution of skill-biased, 183–84; effects on wages and employment of natives, 178–79, 184–88; factors driving decisions and, 182–83; fiscal effects in, 190–91; labor market performance of migrants in, 189–90; length of stay and, 175; net gains from, and option value of waiting, 195–97; points system in, 175, 193–94; quotas and, 175; reasons for existence of, 194–95; restrictions and, 20; wage rigidities and unemployment benefits in, 181–82; welfare and, 192
- Minimum wages, 14, 20, 29–49, 277. *See also* Wages; collective bargaining and, 36; in competitive labor market, 33; cross-country comparisons in, 30–32; defined, 29; in dual labor markets, 37; effects of hikes in, on United States fast-food industry, 40–43; employment protection legislation and, 277, 278; in France, 113; implementing, 15; migration policies and legislation, 176; national government-legislated, 29–30; in noncompetitive labor market, 34–36; positive effects

Minimum wages (*Continued*)

- of, on welfare, 37–38; ratio of, to average wage, 30–31; reasons for existence of, 46; in reducing earnings inequality and poverty, 45–46; studies of: based on firm-level data, 38–39; based on natural experiments, 39–43; based on workers' histories, 43–44; unions and, 76
- Monopoly: bilateral, 36; pure, 34
- Monopsony power, 35; degree of, 35–36; marginal hiring costs and, 34–35; training and, 163, 165; working hours and, 107
- Moral hazard: family policies and, 153; labor market institutions and, 20; unemployment benefits and, 244, 283
- Multivariate exit routes from work, 126–27
- Nash-bargaining rule, 26, 77–79; balanced budget and, 253; unions and, 77, 78; wage bargaining and, 223, 251; wage effects and, 236
- Natives: skill composition of, 184; wages and employment of, 178–79, 184–88
- Negative income tax (NIT), 92–93, 96
- Net replacement rate, 226–29
- Neutrality: employment protection legislation and, 204–6; removing risk, 206
- Nonemployment benefits, 15
- Nonemployment rate, 11
- No-shirking condition, 217
- Oaxaca decomposition, 68
- Objective just cause dismissals, 281–82
- Older workers: experience of, 130; hiring, 129; unemployment of, 129; wage subsidy systems for, 135
- On-the-job training, 157, 163, 169–70. *See also* In-company training
- Optimal job search intensity, unemployment benefits and, 249–50
- Optimal retirement age, 137–38
- Optimal years of schooling, 172–73
- Option value: as theory of migration, 194, 195, 197; as theory of work, 124
- Overtime hours, 102
- Overtime pay, 102
- Parental leave, 142–44; policies on, 139
- Participation effect, 230
- Part-time work, 107–9, 113–16, 144; increase in, 101; reasons for government, 117
- Pay-as-you-go system, 122
- Payroll taxes, 81–100; in competitive labor market, 98–100; composition of, 81; credits or wage subsidies in, 96–97; education and training and, 158; empirical evidence in, 90–92; incidence of, 90–92; measures and cross-country comparisons in, 82–86; negative income taxes or in-work benefits and, 92–96; theory on, 86–90
- Pensions: defined benefit, 122, 124; defined contribution, 122, 125; policies reducing early retirement, 135–36; programs for, 122
- Points system in migration policies, 175, 193–94
- Potential labor force, 13
- Poverty: effectiveness of minimum wage in reducing, 45–46; as trap, 272
- Power couples, residential choices of, 188
- Procedural costs, employment protection legislation and, 199
- Productivity: age and, 132–34; education and, 157, 163, 171, 174, 287; wages and, 130
- Product markets: degree of competition in, 289; employment bias of institutions and, 27–28; institutional reforms and, 21–24
- Progressive tax systems, 82; redistributive purposes of, 87
- Public employment services, 255; need for, 271
- Quotas, migration policies and, 175
- Rent-seeking behavior of unions, 72–73
- Reservation wages, 3, 230, 235; aggregate labor supply and, 8–9; defined, 5; with hours restrictions, 6–7; labor supply and, 4–7; unemployment benefits and, 231–32, 238; with and without hours restrictions, 6
- Retirement. *See* Early retirement programs

- Retirement age: optimal, 137–38; reasons for increasing mandatory, 134–35
- Retirement plans, 15
- Retirement programs, education and, 287–88
- Right-to-manage agreement, 57–58
- Roy model: labor market performance of migrants and, 183, 189, 195
- Schooling, 160–63. *See also* Education; Training; age for compulsory, 15, 171; formal, 157; as investment, 157; optimal years of, 172–73; return to, 165–69
- Self-financed training, 169
- Self Sufficiency Project (SSP), 95
- Severance payments, 204–5; employment protection legislation and, 199
- Skill-based migration, effects on income distribution of, 183–84
- Social pacts, implementing income policies in, 56
- Social security contributions as deferred consumption, 81
- Spike: at benefit expiration, 241; at minimum wage, 39
- Spillover effects: of minimum wage, 47; of unions, 68
- Statistical discrimination, 66
- Subjective just cause dismissals, 281–82
- Subminimum wage, 30
- Subsidized employment, 255
- Substitution effects, 263; fiscal, 263
- Targeted Jobs Tax Credit (TJTC), 97
- Target efficiency, 45
- Tax credits, 97
- Taxes. *See also* Payroll taxes: average rates of, 82; composition of income, 81; employment protection legislation and, 199, 209–10; on labor, 15, 20–21; labor market institutions and, 278–80; negative income, 92–93, 96; progressive systems, 82, 87; unemployment benefits and, 230, 234, 285–86; value-added, 83–84
- Tax rates, 82; average, 82; marginal, 82
- Trade unions. *See* Unions
- Training, 163–65. *See also* Education; Schooling; firm-specific, 157, 163; in-company: government subsidies for, 170; provision of, 158; as investment, 157; market failures concerning, 158; on-the-job, 157, 163, 169–70; paying for general, 174; self-financed, 169
- Transfer component, employment protection legislation and, 199
- Trial costs, employment protection legislation and, 199
- Twins, return to schooling and, 168
- Two-sided job search, 235; effects of employment protection on, 221–24; effects of unemployment benefits on, 249
- Unemployment: bargaining coordination union density and, 71; effect of unemployment benefits on, 226; effect of unemployment benefits on duration of, 239–43; effects of, on generosity of unemployment benefits, 241–42; of older workers, 129; as trap, 272
- Unemployment assistance, 225–26
- Unemployment benefits, 15, 20, 225–53, 277; active labor market policies and, 255, 256, 265, 271, 272, 274, 277, 282–83, 284; balanced budget and, 253; competitive labor markets and, 230–34; defined, 225; dimensions of, 225; duration of, 239–43; effects of unemployment on generosity of, 241–42; efficiency wages and, 236–37; employment protection legislation and, 200, 216, 280–82; entitlement effect and, 237–38; fiscal effects and, 191; generosity index of, 284, 286; imperfect labor market and, 234; indexing, 87, 99; institutions and, 81–82; interaction effect of training and, 261; job search effect and, 234–36; labor market participation and, 251–52; measures and cross-country comparisons, 226–29; migration policies and, 128, 176, 181–82; negative effects of, on job search, 235–36; nonexperimental studies on, 268, 269; numerical simulations and, 253; optimal job search intensity and, 249–50; optimal structure of, 245–47; payroll taxes and, 81–82, 87, 88, 99, 100; as pre-early retirement system, 122; public provision of,

Unemployment benefits (*Continued*)

244–45; reasons for, 247–48; reservation wages and, 231–32, 238; retirement and, 121, 122; shortening duration of, 241–42; taxes and, 285–86, 289; in two-sided job search model and, 249; unions and, 62, 72, 74; vacancies and matching function and, 250; wage bargaining and, 250–51; wage effects of, 236–37; wage rigidities and, 181–82

Unemployment insurance fund, 225

Union density, 52–54; bargaining coordination, unemployment and, 71

Unions, 20; collective bargaining and, 51–79; decline in membership, 53; effect of, on efficiency, 72–74; effects of, on wages, 67–70; employment protection and, 284–85; excess coverage of, 52, 54; free-rider problem in, 62; labor market institutions and, 278–80; migration policies and, 176; power of, in wage bargaining, 17; reasons for existence of, 75–76; regulation of working hours and, 101; strength and efficiency of, 77–79; taxes and, 278–80; wage compression and, 158; wage gaps and, 66, 67

Vacancies: flow value of, 223; higher-paid, 66; for high-productivity jobs, 38; matching of workers and, 4, 37, 65, 221, 222, 235, 236, 250, 259, 260; part-time workers and, 108, 117; posting, 224, 253; public employment services in locating, 271; unemployment and, 2–3, 236, 251, 255, 263, 273

Value-added taxes (VATs), 83–84

Wage bargaining: power of unions in, 17; unemployment benefits and, 250–51

Wage differentials: compensating, 63–65; in imperfect markets, 65–66

Wage distribution, measuring and decomposing effects of unions on, 67–69

Wage effects of unemployment benefits, 234, 236–37

Wage insurance, 45

Wage rigidities, migration policy and, 181–82

Wages. *See also* Minimum wages;

Reservation wages: average production worker, 82–86; bargaining over, 57–58; defined, 2–3; effects of unions on, 67–70; labor demand in determination of, 9–10; ratio of average to minimum, 30–31; relationship between productivity and, 130; reservation, 3; role of unions in affecting setting of, 278–79; unemployment benefits and efficiency, 236–37; union-induced compression of, 158

Wage share, 56

Wage subsidies, 96–97; for older workers, 135

Wedge between labor supply and demand, 26–27

Weekly working hours, regulating, 104–7

Welfare: migration and, 192, 194–95; positive effects of minimum wage on, 37–38

Welfare-to-work schemes, 246

Women: balance of work and family life for, 150–51; child care and, 139, 142; in Norway, 147–48; early retirement programs for, 125–26; employment rates of, 144; fertility and, 149; hours of work for, in Norway, 147–48; maternity leave for, 142; parental leave for, 143–44, 148–49; trade-off between fertility and employment of, 151–53; with young children, 146–47

Work, balance between family life and, 150–51

Workers: activating unemployed, 273–76; discouraged, 12; displaced, 129; employed, 12; inactive, 12, 272; incumbent, 73; older, 129, 130, 135; restrictions on cross-border flows of, 175; unemployed, 12; unemployment among low-productivity, 277

Workfare programs, 272

Work hour regulation, 15, 101–20, 139; collective bargaining and, 101; employment protection legislation and, 101; family policies and, 101; measures and cross-country comparisons in, 101–2; part-time, 107–9, 113–16; reasons for, 116–18;

- reduction of standard, 119–20; theory of, 102–3; unions and, 101; weekly, 104–7; working hours and, 109–13. *See also* Working hours
- Working day, optimal length of, 101
- Working Family Tax Credit (WFTC), 95–96
- Working hours: annual, 110; child care facilities and, 154–55; decline in, per week, 101; mandatory reduction of, in France, 112–13; reduction of standard, 119–20; regulation of, 3, 101–20
- Working week, legal maximum of normal, 102