Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://dare.uva.nl/document/47063

File ID 47063

SOURCE (OR PART OF THE FOLLOWING SOURCE):
Type Dissertation
Title Four essays on education, growth and labour economics
Author M.A. dos Reis Portela
Faculty Faculty of Economics and Business
Year 2007
Pages 121
ISBN 9789051709346

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/221220

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Four Essays on Education, Growth and Labour Economics

ISBN 978 90 5170 934 6
Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul
This book is no. **405** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Four Essays on Education, Growth and Labour Economics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. J.W. Zwemmer
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula
op dinsdag 10 april 2007, te 10:00 uur

door

Miguel Ângelo dos Reis Portela

geboren te Santa Maria da Feira, Portugal

Promotiecommissie

Promotor:	prof. dr. C.N. Teulings
Co-promotor:	prof. dr. R.J.M. Alessie

Overige leden:	prof.dr. G.J. van den Berg
	dr. H.L.F. de Groot
	prof. dr. J. Hartog
	prof. dr. H. Oosterbeek

Faculteit Economie en Bedrijfskunde

Acknowledgements

More than five years have passed since I started my PhD at Tinbergen Institute, and it is now time to remember those who crossed my path towards the current stage. Many people and some institutions have contributed in a non-negligible way to my choices and outcomes.

My promoters, Coen Teulings and Rob Alessie, have provided instructive comments and evaluation of my raw ideas during the thesis process. I specially thank their guidance and their way of challenging my thinking, which made an undoubted contribution to the improvement and completion of this project. In particular, I thank their interaction on Chapter 5: Rob for the thorough discussion on estimation alternatives, and Coen for his valuable comments and questions, which helped clarifying parts of an earlier version of the essay.

The current work benefited from the support of several institutions. In particular, I gratefully acknowledge the financial support by the Portuguese Foundation of Science and Technology (SFRH/BD/5114/2001) and the University of Minho. The essays discussed in Chapters 4 and 5 were only possible thanks to the data set Quadros de Pessoal, made available by the Portuguese Ministry for Labour and Social Solidarity. The help and cooperation of its Department of Studies, Statistics and Planning is highly appreciated. Part of the work on Chapter 5 was carried out while I was visiting IZA, Bonn, whose support I gratefully acknowledge.

Along the road, I collected important comments on different steps of my research. I am thankful to all of those who have somehow contributed to my better understanding of a topic, or who helped me surpassing the difficulties intrinsic to the research work. There are particular contributions that I would like to acknowledge. Chapter 2 embeds valuable comments by Thijs van Rens, Ana Rute Cardoso, and the participants at the Tinbergen Institute Seminar, the 2003 Microeconometrics Conference at CEMFI, the IXth Spring Meeting of Young Economists, and the 2004 European Economic Association and Econometric Society Conference. I must also point out the helpful referee reports on earlier versions of this document. Chapter 3 was born from the insight I received while attending the Empirics of Economic Growth course, taught by Henri de Groot and Raymond Florax. Henri and Raymond made important contributions to a first version of this document. I thank Jacques Silber, Ken Troske, and four anonymous referees for comments on the paper's earlier versions underlying Chapter 4. Finally, I want to thank Dan Hamermesh, Marie-Claire Villeval, and the anonymous referees for valuable comments on the paper

discussed in Chapter 5. I am indebted to Lorenzo Cappellari for having provided the documents on minimum distance estimation that he presented at the 5th UK Meeting of Stata Users.

All the work and effort you go through during a PhD puts at test your endurance. Friends and colleagues at TI have contributed decisively for recharging my batteries. It is with a kind smile, that I will always remember our triple tours across Amsterdam, the dinners and parties organized among us, and the trips we made across The Netherlands. I thoughtfully remind Robert, Sebi, Wendy, Maria, Ana, Matthijs, Aico and Annette, with whom I shared the office at TI, and Ernesto, Aljaz, Suncica, Hugo, Martijn, José, Mauro, Marcos, Simonetta, and Peter Gautier, and all the people who organized, or simply participated, in all the events that brought to us some of the sun that was shinning outside. I am particularly thankful for the dinners organized by Ana and Ernesto, who always found a way to please us with different flavors. My squash time with Nick, Naomi, Hugo and António was equally important to keep my stress at bearable levels. With Viktória, I shared the pleasure of conversation and our interest on photography. António has been a Portuguese friend who accompanied me in this process with lively discussions on work, Portugal, and life in general. Carlos, who became part of our community at a latter stage, conveyed an important contribution to my life experience. Positive interaction among people, both in celebration and at work, has been crucial to my achievements while studying at TI. For this, I thank all the colleagues, who have been around for this period, and Maarten Lindeboom, who inspired this atmosphere since the beginning. Finally, I thank the staff at TI, Marian, Nora, Walther, and Arianne.

Going back to the very beginning of this process, I am indebted to Paulo Guimarães for his key role in creating conditions and pressing young people like me, working at the University of Minho, to pursue their graduated studies in recognized international institutions. Ana Rute Cardoso, as my Master thesis' supervisor, introduced me to scientific research and guided my first steps in economic analysis. Both, Paulo and Ana, molded my goals with their examples. In the final stage of the thesis' preparation, Fernando, my friend and colleague at the University of Minho, has been critical remembering my goals. As Head of Department, he has put the right pressure, while as a friend he gave me the needed support. I am thankful for the friendly dinners with his family after long days of work.

My friends in Portugal took care of the link to my roots. I remember with a warm feeling the Sunday coffees during my short breaks in Portugal. A special thought goes to Óscar, Zé Mário, Sérgio, and Francisco, and all of those that gather around us to listen the stories

from Amsterdam. Astride, Luís, Jorge, and Carla are part of my Portuguese backbone, and reminded me how important it is to preserve one's good spirits.

A smile to my family to overcome the deep tear, that strokes us while I was working on this project. A smile to you, father. Although I keep an empty space inside me, following an unfinished conversation, I have been filling it since with the memories you left. A hug to my uncles, Telmo and Quim, for their care and support.

I dedicate this thesis to Carla, the special person who I was lucky enough to find. She was the one who kept the loose ends tied in all the difficult moments we surpassed. To her, I owe the traces of harmony I am rebuilding today. A ti dedico o sorriso mais singelo e sentido que consigo reunir.

Miguel Portela
Braga, 23 February, 2007

Contents

List of Tables **xiii**

List of Figures **xv**

1 Introduction **1**

2 Measurement error in education and growth regressions **5**

2.1 Introduction 5

2.2 Sources of data on education 7

2.3 How systematic is the difference between census and non-census data? . . . 9

2.3.1 Origins and identification of the systematic difference 9

2.3.2 Data description 11

2.3.3 Empirical evidence 11

2.3.4 How to correct for the systematic difference? 14

2.4 Growth regressions: what changes? 16

2.4.1 OLS estimations 16

2.4.2 Sensitivity analysis: the dynamic model 22

2.5 Final remarks 25

Appendix to Chapter 2 **27**

2.A Additional results and census example 27

3 Economic cross-country spillovers of education **29**

3.1 Introduction 29

3.2 How to measure the interaction between countries? 32

3.2.1 The model 32

3.2.2 Measuring proximity 33

3.2.3 The regression model 36

3.3 Data 37

3.4 Empirical analysis . . . 40
3.4.1 Estimating the proximity measure . . . 40
3.4.2 Testing for the presence of country effects . . . 41
3.4.3 Education spillovers . . . 49
3.5 Concluding remarks . . . 53

Appendices to Chapter 3 **57**
3.A Computation of returns to education . . . 57
3.B Data description . . . 58

4 Gender segregation and the wage gap in Portugal: an analysis at the establishment level **61**
4.1 Introduction . . . 61
4.2 Data set . . . 62
4.3 The Portuguese labour market and trends in female employment . . . 63
4.4 Gender segregation at the establishment level: systematic and random components . . . 65
4.5 The impact of gender segregation on wages . . . 68
4.5.1 Higher concentration of women in the establishment: lower wages for women, but higher for men . . . 70
4.5.2 Segregation and the wage gap . . . 72
4.6 Conclusion . . . 75

Appendix to Chapter 4 **77**
4.A Additional tables . . . 77

5 Micro foundations for wage flexibility: wage insurance at the firm level **85**
5.1 Introduction . . . 85
5.2 Wage insurance in the previous literature . . . 86
5.3 Wage setting institutions in Portugal . . . 88
5.4 Data set . . . 89
5.5 Firm performance . . . 90
5.6 Worker earnings . . . 92
5.7 Insurance provision by the firm . . . 95
5.8 Forces shaping wage flexibility at the firm level . . . 98
5.9 Conclusion . . . 101

Appendices to Chapter 5 **103**
5.A Checks on the consistency of data 103
5.B Constraints imposed 103
5.C Descriptive statistics 104

6 Conclusion **107**

References **113**

Samenvatting (Summary in Dutch) **123**

List of Tables

2.1 Summary statistics and the distribution of census data for 5-year interval data . . . 12
2.2 Education regressions . . . 13
2.3 Correlations among education measures in levels . . . 14
2.4 Correlations among education measures in first-differences . . . 15
2.5 The effect of education on growth - annualised OLS estimations . . . 17
2.6 Growth regressions - comparison with Cohen and Soto data . . . 22
2.7 Dynamic income regressions – 10 year data . . . 23
2.8 Dynamic income regressions – 5 year data . . . 27
2.9 Example of census variables for Figure **??** . . . 27

3.1 Summary statistics . . . 39
3.2 Log average trade regression . . . 40
3.3 Estimated country effects *vs* education . . . 41
3.4 Testing for the presence of country effects, specification I . . . 43
3.5 Testing for the presence of country effects, specification II . . . 44
3.6 Testing for the presence of country effects, specification III . . . 45
3.7 Testing for the presence of country effects, specification IV . . . 46
3.8 Income regressions . . . 50
3.9 Specific returns to education . . . 52
3.10 List with the 75 countries used in the analysis . . . 58
3.11 Sample of weights for The Netherlands . . . 59
3.12 Sample of weights for The United States . . . 59

4.1 Gender segregation at the establishment level . . . 68
4.2 Male/female log-wage decompositions . . . 73
4.3 Contributions to the observed gender wage gap (%) . . . 74
4.4 Sample mean values (males) . . . 77

4.5 Sample mean values (females) 78
4.6 Ordinary least squares regressions (males) 79
4.7 Ordinary least squares regressions (females) 80
4.8 Firm fixed-effects regressions (males) 81
4.9 Firm fixed-effects regressions (females) 82
4.10 Contribution of the share of females to the total gender wage gap, under alternative estimation methods 83

5.1 Sales regression 91
5.2 Firms' autocovariances 92
5.3 Wage regression 93
5.4 Workers' autocovariances 94
5.5 Testing for insurance 96
5.6 Insurance heterogeneity 99
5.7 Descriptive statistics on workers 104
5.8 Descriptive statistics on firms 106

List of Figures

2.1 Plot of education with census and non-census data 10
2.2 Education information for Argentina 16
2.3 Returns to education for the different time spans and education variables . 21

Chapter 1

Introduction

This thesis is organized into four chapters, each corresponding to one paper. As a result, each chapter constitutes an autonomous piece of work, self–contained in its motivation, methods, and results. This introduction is therefore kept brief, leaving a more extensive motivation for each of the chapters. Throughout the different chapters the approach is mainly an empirical one, embracing both macro- and microeconomic perspectives.

The first part of the thesis, comprising Chapters 2 and 3, concentrates on the role of education in economic growth, and the analysis is performed using cross-country data. The underlying question is whether education has a positive effect on countries' economic performance. The issues of measurement error in education and education spillovers are covered in Chapters 2 and 3, respectively. In the second part of the thesis, the focus switches to the microeconomic analysis of the labour market, using a longitudinal matched employer-employee data set for Portugal. Chapter 4 analyses gender segregation and the wage gap, and Chapter 5 deals with wage flexibility and the provision of wage insurance by firms. A brief outline of each of the four chapters follows.

Chapter 2 is based on Portela et al. (2004), and discusses the issue of measurement error in countries' education and its effect on growth regressions. This issue is recognized as an important source of bias in growth regressions, and as such it plays a relevant role in the empirics of economic growth. The perpetual inventory method used by Barro and Lee (2001) for measuring countries' education level leads to systematic measurement error. We show that there is a systematic difference in the education level between census data and observations constructed from enrolment data, and discuss a methodology for correcting the measurement error. The major contribution of our analysis is thus to typify the structure of the measurement error in one of the most used sources of information on cross-country education, and to propose a simple procedure for its correction.

In this chapter we also analyse the effect of such measurement error on GDP regressions. We compare our estimates with those of Topel (1999) and Krueger and Lindahl (2001), and replicate the analysis with the data provided by Cohen and Soto (2001).

The standard attenuation bias suggests that using the corrected data would lead to a higher impact of education on economic growth. However, our estimation results reveal the opposite. We then discuss why the measurement error yields an overestimation, and provide an explanation for the difference in results between growth regressions based on 5 and on 10 year first–differences. In doing so, we consider both immediate and long run returns to education associated with the different time spans of the data (5 and 10 years, respectively), as well as the underlying half-life time. As a robustness check, we estimate a dynamic panel data version of the simplified growth model.

In Chapter 3, we analyse cross-country education spillover effects. By focusing on country data, our goal is to measure global economic externalities of education. Our analysis fits into the broader discussion of the existence of social returns to education. Both policymakers and economists have particular interest in the topic because of, among other aspects, its implications for economic growth and income inequality across countries. We propose a measure of neighbouring education, which is used to re-evaluate the role of education in GDP regressions. Our goal is twofold. First, an unambiguous identification of human capital externalities is of major interest for the economic growth literature. Second, by using a high level of aggregation we contribute with a different perspective to the analysis of education externalities, and overcome some difficulties that are common to studies that use more disaggregated data.

We start by discussing the issue of proximity between countries, and introducing the definition of neighbourhood to be used in the remainder of the chapter. Proximity is based on economic distance and translates the interaction between countries. We develop a procedure for measuring proximity, which is based on countries' bilateral trade. We then estimate a dynamic specification of the economic growth regression model, which includes both country's own and neighbouring education, as well as country effects. We test for the presence of such effects, with results showing that country effects are absent from the main specification. Based on this result, we estimate the main economic growth regressions by means of nonlinear least squares. Such solution allows for the retrieval of the estimate of a proximity decay parameter associated with the definition of neighbourhood. In order to illustrate the results, we discuss the returns to education for three individual countries: The Netherlands, The United Kingdom and The United States. Finally, we extend the analysis of the returns to education performed in Chapter 2, by making the distinction between internal, external and total returns to education.

In the following two chapters the analysis is performed at the example of the Portuguese labour market. The peculiarities of its labour market and the unique features of the available matched employer-employee data, make Portugal a very appealing case

study. After joining the European Union in 1986, Portugal experienced a period of intense modernization and restructuring. The good economic performance that characterized the first phase of that process was followed by a recession period, which tended to last longer than for most EU countries. Over that period the intense demand for skilled workers, contrasted with an average low level of education on the supply side. The returns to education have increased, and wage inequality remained quite high. A high wage flexibility is combined with a high level of employment rigidity. Finally, the female participation rate is high, and has increased over that period, at the same time that an improvement in the skill composition of the female labour force was taking place. Regarding the data set, it is important to highlight the high rate of response and the low incidence of measurement error. The data collection results from an inquiry. It resembles a census, since all firms in the private sector with wage-earners have to fill it in, and provides full coverage of the paid labour force within each firm. The data include extensive information on firms and its employees, and has a high degree of reliability, reinforced by legal requirements.

In Chapter 4, which resembles Vieira et al. (2005),[1] we analyse the trend in worker segregation at the establishment level and its impact on wages in Portugal over a fifteen year period. We concentrate on the gender dimension to answer the following three questions: (i) what is the level of gender segregation across establishments in the Portuguese labour market and how has it evolved over time?; (ii) what is the impact of segregation on wages?; and (iii), is that impact different for men and women? In order to account for biases in the computation of traditional segregation measures, we quantify the degree of systematic segregation as a departure of overall segregation from random segregation. We use standard wage decomposition techniques to evaluate the impact of the composition of the labour force at the establishment level on wages. The issue of gender segregation across establishments is of major interest for policymakers. On the one hand, employment segregation is relevant in the context of wage inequality since it is a source of wage differences between groups. On the other hand, characterizing sources of wage differences by gender is of interest to better tackle women discrimination in the labour market.

Chapter 5, a revised version of Cardoso and Portela (2005), analyses the provision of wage insurance by the firm, in the context of microfoundations for wage flexibility. Wage flexibility is often pointed out in the literature as a mechanism that can reduce macroeconomic fluctuations and improve macroeconomic performance. The basic reasoning is that, once negative shocks happen, wage flexibility will enable the price of labour to adjust. Instead, if wages were rigid, the adjustment would be promoted via the "quantity" of labour,

[1] The original publication is available at www.springerlink.com: http://www.springerlink.com/content/l617j50071766204/.

leading to higher unemployment. There are two aspects to wage flexibility, although only one has deserved considerable attention in the literature. First, wage flexibility can be evaluated as the degree of responsiveness of wages to aggregate conditions, in particular the unemployment rate. Second, wage flexibility can be assessed as the responsiveness of wages to changes in the firm-level conditions, for example, the firm productivity or the demand for the firm ouput. The latter issue has been less addressed mainly because of lack of adequate data.

We will concentrate precisely on that aspect, which can be understood as the micro foundations for wage flexibility. Specifically, we evaluate the impact of product market uncertainty on wages, addressing the questions: (i) what is the responsiveness of wages to shocks to firm output?; (ii) which firm and worker attributes are associated with a higher degree of wage flexibility at the micro level? We check in particular the role of regulations in the labour market constraining the responsiveness of wages to firm shocks. We first rely on Guiso et al. (2005), estimating dynamic models of sales and wages to evaluate the sensitivity of wages to permanent and transitory shocks to firm performance, and then explore the factors associated with higher wage flexibility.

Finally, the main results presented in these four essays are summed up in Chapter 6. Some policy implications, as well as lines for future research, are addressed.

Chapter 2

Measurement error in education and growth regressions

2.1 Introduction

Measurement error in education is widely recognized as an important source of bias in growth regressions; see for example Krueger and Lindahl (2001). This chapter shows that the way Barro and Lee (2001) constructed the education data yields a systematic error. Some data points are directly derived from census observations, others are derived from the previous census information, using enrolment data and the perpetual inventory method for updating. We show that this updating yields a systematic measurement error, as it yields an underestimation of the growth of education during the period. Previous attempts to correct for this error have either only been successful for a limited number of countries or were based on arbitrary corrections made by the researcher (see De la Fuente and Doménech, 2002, and Cohen and Soto, 2001). Our analysis leads to a simple correction procedure for data points based on the perpetual inventory method that does not require any *ad hoc* decisions.

The issue of the measurement error in education data is of great practical relevance for the interpretation of the relation between education and GDP. There are two main approaches to model human capital in the context of economic growth: (i) Nelson and Phelps (1966), and (ii) Lucas (1988). In the former, human capital is crucial to innovate and adopt new technologies. Hence, the *growth* rate of output is determined by the *level* of human capital. In the latter, human capital is interpreted as a normal input in the production process. Hence, *changes* in output are determined by *changes* in the human capital stock. The estimated effect of education on economic growth depends on the reliability of education data. Benhabib and Spiegel (1994) and Barro and Sala-i-Martin

(1999) conclude that it is the level of education, not its change, that has an impact on economic growth, which is evidence in favour of Nelson and Phelps' argument that growth is driven by the stock of human capital. Krueger and Lindahl (2001) argue that these conclusions are highly affected by measurement error in the average education of countries. The problems of measurement error are exacerbated when taking first differences. First differences reduce the signal and increase the noise. Hence, the signal-noise ratio falls dramatically by first differencing. Krueger and Lindahl's solution to this problem is to increase the differencing period from 5 to 10, or 20, years, thereby increasing the signal. They show that indeed the coefficient on the change in education increases by taking a longer differencing period. The authors conclude that "the change in education is positively associated with economic growth once measurement error in education is accounted for" (Krueger and Lindahl, 2001, p.1130), finding empirical evidence in favour of Lucas' argument. However, the problem with this conclusion is that the Nelson and Phelps (1966) model would lead to exactly the same conclusion. When the *level* of education affects output *growth*, then the effect of the *level* of education on output increases linearly with the differencing period, and hence, so does the effect of the *change* in education.

From our analysis, first we conclude that measurement error in education data is important. We find large and statistically significant differences between data points based directly on census information and data updated with the perpetual inventory method. One would expect that these differences have large effect on growth regressions, in particular where differencing exacerbates the problem, in particular when using 5 year differences. Many countries hold a census every ten years, so that 5 year differences switch back and forth between direct census information and updating by the perpetual inventory method. This turns out not to be the case. Using our corrected measure of education reduces the coefficient on changes in education. This runs counter to the standard argument of contamination bias, which is supposed to lead to lower coefficients when using data spoiled by measurement error. The reason for this paradox is that, in the standard model, measurement error increases the variance of the explanatory variable, since the measurement error is supposed to be orthogonal to the signal. In this case, the measurement error decreases the variance, since the perpetual inventory method smoothens observations at the beginning and the end of the observation period, thereby compressing the data. However, our exercise contributes to the explanation of the differences in the coefficient of the change in education based on a 5 and a 10 year differencing period, and compares to Krueger and Lindahl (2001). All this leads to the inevitable conclusion, previously obtained by Teulings and van Rens (2006), that education has a moderate immediate effect on GDP of about $4.2 - 6.5\%$, but a huge long-run effect of about $54 - 59\%$, which however takes

ages to materialize, the half value time being $75 - 99$ years. This conclusion is obviously conditional on the identifying assumption that is used in whole this literature till so far, that current innovations in GDP have no effect on current innovations in education.

The chapter is organized as follows. In the next section we describe different sources of data on education. Then, in Section 2.3 we will analyse how systematic is the difference between census and non-census data. In Section 2.4 we will concentrate on the interaction between education and growth using the corrected data on education, comparing the results with known figures. Finally we conclude.

2.2 Sources of data on education

The most used data set on international education attainment is the one released by Barro and Lee (2001).[1] They build their data on educational attainment from census or survey data. When this information is not available, the authors use a perpetual inventory method based on enrolment data in order to generate either a forward-flow, or a backward-flow. The flows are constructed from the benchmark stocks defined by the census or survey data. For intermediate observations, the constructed data point is a weighted average of the forward-flow and the interpolation between two benchmarks. For the observations before the first and after the last census or survey, interpolation is infeasible. Then, the constructed data apply by either the forward- or the backward-flow to the closest available census or survey data point. The enrolment data are adjusted for repeaters and changes in the duration of years of schooling.

Barro and Lee's data received criticism. De la Fuente and Doménech (2002, p.1) construct a revised version of the Barro and Lee (1996) data set for a sample of 21 OECD countries "using previously unexploited sources [and following] a heuristic approach to obtain plausible time profiles for attainment levels by removing sharp breaks in the data that seem to reflect changes in classification criteria". The authors state that "to avoid unreasonable jumps in the series [they proceed] by choosing the most plausible figure when several are available for the same year, and by reinterpreting some of the data (...) when it

[1] Alternative sources are Kyriacou (1991) and Nehru et al. (1995). The latter ignores census data. De la Fuente and Doménech (2002, p.6) criticise this choice, and argue that it is difficult to justify "discarding the only direct information available on the variables of interest."

seems sensible to do so" (De la Fuente and Doménech, 2002, p.13).[2] Missing observations are filled in, if possible by interpolation, or otherwise by back- and forward projections. The authors "avoided the use of flow estimates based on enrollment data because they seem to produce implausible time profiles" (De la Fuente and Doménech, 2002, p.14). The authors state that "the construction of our series involves a fair amount of guesswork," and that their data "look more plausible than most existing series, at least in terms of their time profile" (De la Fuente and Doménech, 2002, p.14).

Cohen and Soto (2001) extend the work of De la Fuente and Doménech (2002) to several other countries. An important difference to De la Fuente and Doménech is that Cohen and Soto allow for the use of enrolment data when needed. The authors have constructed a data set for 95 countries with information on education achievement from 1960 to 2000, for ten year intervals, plus a projection for 2010. Their methodology is to "minimize the extrapolations and keep the data as close as possible to those directly available from national censuses" (Cohen and Soto, 2001, p.6). They argue that some of the differences between their data and those provided by Barro and Lee (2001) can be explained by: (i) divergences in classification; (ii) the use of more census information than Barro and Lee; (iii) the use of a different methodology for extrapolating the missing data; (iv) errors in Barro and Lee data.

The conclusion is that, in spite of the improvements in data, so far measurement error in education data remains a problem. The Barro and Lee data are highly erratic. For example, in many cases, the average education level decreases over time within countries, which does not fit casual observation. De la Fuente and Doménech's data is a valuable effort, but requires a large amount of ad hoc decisions and is only available for a sample of 21 countries. Cohen and Soto's data increases the sample size, but is only available on 10-year intervals. Also, both these data sets face the criticism that measurement error problems were not entirely solved. Finally, Kyriacou data is very problematic given the estimation procedure used,[3] and the fact that it is only available for the period 1965–1985.

[2] These two data sets are not directly comparable since Barro and Lee's data is based on people having completed some educational level, while De la Fuente and Doménech's data apply to people who have attended some educational level. The study by De la Fuente and Doménech (2002) has recently been published as De la Fuente and Doménech (2006). However, the data set is the same, and they refer to the working paper for data details.

[3] Kyriacou assumes that the relationship between average years of schooling in the labour force and the enrolment ratios in primary, secondary and higher education is relatively constant over time and across countries.

2.3 How systematic is the difference between census and non-census data?

2.3.1 Origins and identification of the systematic difference

The hypothesis we will test is that the methodology used by Barro and Lee to impute missing values in the Barro and Lee data underestimates the true values of education. This underestimation results from the assumption that people's survival rate is independent of the educational level. In their own words, Barro and Lee (1993, p.374) state that "some error is introduced (...) if educational attainment is growing rapidly, because the older people then have less human capital and a greater probability of dying." If average education within a country is rising, as it seems to be the case for an important portion of the countries, the implication would be an underestimation of the educational attainment. The increase in the schooling level of a population occurs mainly because the younger generations are more educated. In this case, the estimation procedure underestimates the survival of more educated individuals, resulting in a lower attainment for the country as a whole. The same idea is identified in Barro and Lee (2001, p.545), when they say that "in a typical country in which educational attainment is growing, mortality would be higher for the older people who are less educated. Then the assumption of uniform mortality can cause a downward bias in the estimation of the total educational stock."

If this is true, we should observe in the data that: (i) the increase in education between two consecutive census observations should be higher than the increase between non-census observations, and (ii) the education level jumps upward between a non-census to a census observation, and that this jump is larger, the larger the period since the previous census. Figure 2.1 shows the argument for a hypothetical country with 9 observations. At the horizontal axis we have the periods, while the vertical axis plots the average education level in each period. The steeper and darker line represents the evolution of true education. For simplicity, we assume that true education follows a constant trend. The observations represented by an empty square, located on this line, represent the census information available. The circle dots represent the estimated points using the enrolment data and the benchmark census information. We also assume that the estimation process leads to a constant trend, which underestimates the true value. The lighter line represents this. The filled square dots represent the values of education that would be estimated for periods in which we have census data.

The change in education from period 4 (the empty circle in period 4) to period 5 (the empty square in period 5) can be decomposed as the variation predicted by the

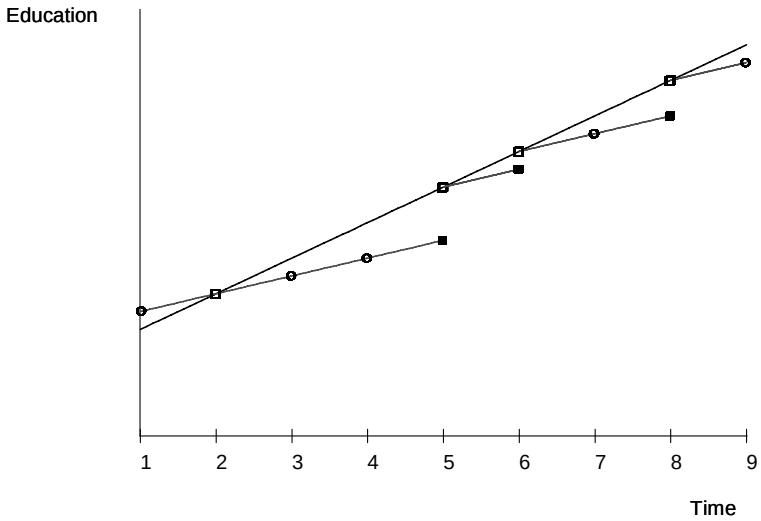

Figure 2.1: Plot of education with census and non-census data.
Notes: This plot refers to a hypothetical country. The empty square is used for census information; the circle dots represent the estimated points using the enrolment data and the benchmark census information; and the filled square dots represent the values of education that would be estimated for periods in which we have census data.

perpetual inventory method (the filled square dot over the lighter line in period 5), plus the accumulated errors since period 2, originated by the underestimation. The jump between the non-census (hypothetical) and census data points in period 5 (the difference between the filled square dot and the empty square dot) is proportional to the time elapsed since the previous census. In period 3 the error is given by the distance between the empty circle and the steeper line. In period 4 the difference between the empty circle and the steeper line gives the accumulated error in period 3 and 4. The error specific to period 4 can be retrieved if we imagine a non-census trend line departing from the true education value in period 3.

Barro and Lee's procedure implies that missing values are constructed differently according to the type of observation: (i) observations before the first census; (ii) observations between two census observations; (iii) observations after the last census. Our empirical strategy tests for systematic differences between census and non-census observations, where we take into account the differences between these three types. We constructed four variables. $Before$ applies to type (i) observations; it measures the time interval till the first census. $Last$ and $LastC$ apply to type (ii) observations; the first records the number of periods since the previous census, while the second also records the lag till the previous census, but just for census data points, being zero otherwise. $After$ applies to type (iii)

and measures the lag till the last census.[4] These variables adequately cover the hypothesised bias introduced by Barro and Lee's procedure as depicted in Figure 2.1. If we used just a dummy for non-census observation instead, then its coefficient would be a weighted average of the changes associated with different lags till the previous census. Moreover, it would not have differentiated between the positive bias for type (i) observations and the negative bias for type (ii) and (iii) observations.

2.3.2 Data description

Table 2.1 provides a description of the data. We will focus our attention on population aged 15 and over. The dummy variable $Census$ assumes the value 1 for observations based on a census or survey, and 0 otherwise. The variables $Before$, $Last$, $LastC$, and $After$ are constructed from $Census$ variable as described above. The income variable is real Log GDP per worker, $LGDP$, and is obtained from the Penn World Table 6.1 (Heston et al., 2002). All variables are available on five year intervals, between 1960 and 2000. Average income increased by 18% per decade, while average education increased by 0.70 years, achieving 6.33 years in 2000. Its dispersion has been relatively stable over time, with a slight increase in the beginning of the sample period. With the exception of 1985, income dispersion has steadily increased. Only 32% of the information on education is based on census/survey data. There is a concentration of census information at the start of each decade, 1970, 1980, and 1990. This is a particularly relevant feature when first differencing the data using a 5 year time frame. 46% of the countries have 2 or fewer census observations, and only 26% have 4 or more. Finally, the distribution of countries per period is relatively balanced.

2.3.3 Empirical evidence

Consider the following model:

$$Edu_{it} = \gamma_t + \beta_1 Before_{it} + \beta_2 Last_{it} + \beta_3 LastC_{it} + \beta_4 After_{it} + \eta_i + \varepsilon_{it} \tag{2.1}$$

where Edu_{it} is the education level of country i, in period t; γ_t is the specific effect for period t; η_i is country i's specific effect; and ε_{it} is a white noise error term. Taking

[4] See Table 2.9, in Appendix 2.A, for an example of these variables corresponding to Figure 2.1.

Table 2.1: Summary statistics and the distribution of census data for 5-year interval data

Variable	Statistic	1965	1970	1975	1980	1985	1990	1995	2000	Total
Observations		104	106	110	111	112	116	111	111	985
Census	Mean	0.20	0.54	0.35	0.60	0.20	0.39	0.10	0.00	0.32
Before	Mean	0.61	0.30	0.12	0.02	0.00	0.00	0.00	0.00	0.23
	% zeros	66	81	90	98	100	100	100	100	88
Last	Mean	0.46	1.04	1.03	1.29	0.85	0.91	0.17	0.00	0.64
	% zeros	54	39	32	32	49	61	90	100	62
LastC	Mean	0.03	0.68	0.42	0.95	0.30	0.84	0.17	0.00	0.38
	% zeros	97	62	76	49	83	65	90	100	80
After	Mean	0.03	0.07	0.16	0.37	0.83	1.38	2.27	3.27	0.95
	% zeros	97	96	90	79	54	42	10	0	62
Education	Mean	3.90	4.28	4.52	4.99	5.31	5.84	6.07	6.33	5.03
	Std.Dev.	2.56	2.70	2.75	2.86	2.80	2.84	2.80	2.82	2.88
ΔEdu	Mean	0.13	0.44	0.35	0.52	0.32	0.47	0.32	0.26	0.35
	Std.Dev.	0.29	0.57	0.40	0.60	0.35	0.53	0.32	0.14	0.44
Observations		104	104	106	110	111	112	111	111	869
LGDP	Mean	8.98	9.09	9.16	9.24	9.27	9.33	9.39	9.50	9.20
	Std.Dev.	0.95	0.99	1.00	1.04	1.02	1.07	1.11	1.13	1.04
Observations		85	89	92	94	95	97	97	89	821
$\Delta LGDP$	Mean	0.17	0.16	0.11	0.10	0.04	0.04	0.05	0.08	0.09
	Std.Dev.	0.12	0.14	0.15	0.17	0.17	0.17	0.18	0.10	0.16
Observations		83	85	89	92	94	95	97	87	722
Distribution of census data										
Number of census		1	2	3	4	5	6	7	8	
Number of countries		25	28	33	22	6	0	1	1	116

Note: The summary statistics for ΔEdu and $\Delta LGDP$ are for changes over five year periods.

first-differences eliminates the fixed country effect:

$$\Delta Edu_{it} = \Gamma_t + \beta_1 \Delta Before_{it} + \beta_2 \Delta Last_{it} + \beta_3 \Delta LastC_{it} + \beta_4 \Delta After_{it} + \Delta\varepsilon_{it} \quad (2.2)$$

where Δ is the first difference operator, and Γ_t are period specific effects. Estimation results are presented in Table 2.2.[5] In column 1 we report the estimation of equation (2.1), using the fixed-effects estimator. Columns 2 and 3 report the results for equation (2.2), where we use OLS in column 2 and fixed effects in column 3. For the model in levels in column 1 the hypothesis of the absence of country specific effects is rejected. We also reject the hypothesis that the level error terms are not serially correlated. When we

[5] The standard errors are robust to heteroskedasticity and error correlation within countries.

Table 2.2: Education regressions

	Levels	First-differences	
Variable	Fixed Effects	OLS	Fixed Effects
Before	0.391**	0.250**	0.140*
	(0.073)	(0.055)	(0.066)
Last	-0.200**	-0.198**	-0.186**
	(0.032)	(0.028)	(0.031)
LastC	0.199**	0.202**	0.193**
	(0.033)	(0.029)	(0.029)
After	-0.214**	-0.272**	-0.316**
	(0.057)	(0.057)	(0.086)
Wald: census variables	20.322**	17.176**	14.076**
Wald: time dummies	74.741**	3.353**	5.558**
F-Test	251.61**		1.09
AR(1)	159.194**	-1.070	0.000
Observations	985	869	869
Countries	116	116	116

Notes: Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. Robust standard errors in parentheses. In the first column the dependent variable is Edu, while in the following two columns it is ΔEdu. All regressions include time effects. Wald stands for the Wald test on joint significance of the refered set of variables. F-Test reports the F statistic for the test of absence of country specific effects. For the first column the null is stated as 'H_0: all $\eta_i = 0$'. AR(1) is the test for first order serial correlation in the errors.

apply the fixed-effects estimator to the first-differences specification (column 3) we do not reject the hypothesis that all the country specific effects are equal to zero. Hence, the first differences without fixed effects (column 2) is the preferred estimation. The subsequent discussion is restricted to this model.

The estimation results strongly confirm our hypothesis regarding the biases in non-census observations. All four variables have the expected sign and are highly significant. The coefficients on $Last$ and $LastC$ are identical in absolute value, as predicted. Furthermore, the coefficient on $After$ and $Before$ are larger in absolute value than the coefficient on $Last$. This too fits our hypothesis. Since the type (ii) observations are a weighted average of interpolation between neighbouring census observations and a forward perpetual inventory method, while type (i) and (iii) are fully based on the perpetual inventory method, the bias is larger for the latter group of observations. The magnitude of the bias is huge, some 0.20 year per 5 year period, or about 60% of the total average increase of education per 5 year period. The fill in procedure of the observations for which no census information is available introduces therefore a large and systematic bias in the data.

Given the fact that many countries hold a census every 10 years (usually at the beginning of a decade), the systematic bias in the non-census observations yields a particular erratic time series of first differences when using a 5 year period.

2.3.4 How to correct for the systematic difference?

How can we use this information to improve the quality of the data? Our idea is to use the regression results to correct the original data by the subsequent expression

$$PEdu_{it} = Edu_{it} - \hat{\beta}_1 Before_{it} - \hat{\beta}_2 Last_{it} - \hat{\beta}_3 LastC_{it} - \hat{\beta}_4 After_{it} \tag{2.3}$$

where $PEdu_{it}$ is the corrected education variable.[6] The estimated coefficients are the ones reported in column 2 of Table 2.2.[7]

Table 2.3: Correlations among education measures in levels

	Edu	PEdu	EduCS	EduDD	Mean	Variance
Edu	1				5.028	8.299
	(985)				(985)	(985)
PEdu	0.987	1			5.281	8.883
	(985)	(985)			(985)	(985)
EduCS	0.956	0.956	1		5.683	9.957
	(420)	(420)	(420)		(420)	(420)
EduDD	0.892	0.888	0.933	1	9.567	4.464
	(155)	(155)	(80)	(155)	(155)	(155)

Notes: The reported numbers are correlations between pairs of the four education variables. Number of observations in parentheses. The last two columns report the mean and the variance of each variable.

Tables 2.3 and 2.4 give the correlations between the various education variables; Table 2.3 in levels and Table 2.4 in first-differences. The correlation between Barro and Lee education level and the corrected education variable is high, 0.99. The correlation between these two variables and the series constructed by Cohen and Soto (2001) ($EduCS$) and De la Fuente and Doménech (2002) ($EduDD$) is only slightly lower. The mean of Barro and Lee data is the lowest of all four, while $EduDD$ presents the lowest variance. For

[6] With this formulation we impose the same bias correction across countries. Although we acknowledge that this is not the most realistic assumption, sample size limitations restrict the available alternatives to implement corrections specific to countries.

[7] The data used in this analysis is available at http://www.tinbergen.nl/~portela/education.

Table 2.4: Correlations among education measures in first-differences

	DEdu	DPEdu	DEduCS	DEduDD	Mean	Variance
DEdu	1				0.350	0.192
	(869)				(869)	(869)
DPEdu	0.888	1			0.501	0.147
	(869)	(869)			(869)	(869)
DEduCS	0.369	0.348	1		0.843	0.187
	(335)	(335)	(335)		(335)	(335)
DEduDD	0.068	0.019	0.391	1	0.376	0.020
	(135)	(135)	(60)	(135)	(135)	(135)

Notes: Number of observations in parentheses. For *DEduCS* first-differences are computed over a 10 year interval. Correlations with the other variables account for this adjustment. See note to Table 2.3.

the data by De la Fuente and Doménech, this comparison does not make much sense, since they consider only the very selective sample of 21 OECD countries. To a lesser extent, a similar objection can be raised against a comparison to the Cohen and Soto data, where the difference in the number of observations is mainly due to the fact that they have data once every 10 years. However, the comparison with our corrected data is highly informative. The bias in the fill in procedure in the Barro and Lee data leads to an underestimation of the average education level by 0.25 year. Even more importantly, it leads to an underestimation. An eye on Figure 2.1 reveals why this is the case. The bias understates the final observations, but overstates the initial observations, leading to a compression of the "true" variance. So contrary to the classical model, where measurement error is orthogonal to the signal and therefore increases the variance of the observed data, here the measurement error compresses the variance.

In first differences, the correlations between education variables are much lower. The correlation between Barro and Lee and our corrected variable is still high, 0.89. For alternative sources of information, the correlations drop significantly. Once more, the correlations are higher with Cohen and Soto's data. Again a comparison of the mean and variance of the changes between Barro and Lee and our corrected variable is revealing. The bias in Barro and Lee compresses the measured average growth of education substantially, from 1.00 year per decade to 0.70 year. The variance of the changes is however overestimated in the Barro and Lee data, as one would expect with all the erroneous changes back and forth from census to non-census based observations.

These ideas are well documented by the data on Argentina, as shown in Figure 2.2. We observe spikes at each census observation for the data estimated by Barro and Lee

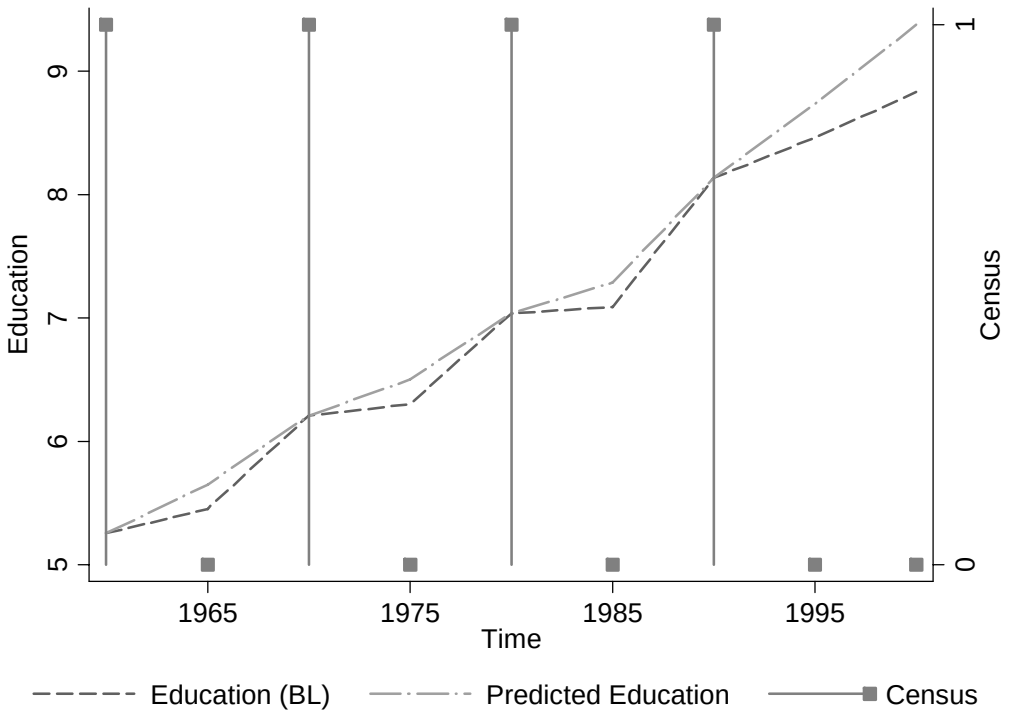

Figure 2.2: Education information for Argentina

(Education (BL)). Between census (1960–1990), our procedure (Predicted Education) smoothens the data. However, for observations after the last census available (1990), the constant correction induces a higher variance. When the variables are analysed in changes, $PEdu$ has a higher mean, but a smaller variance than Edu. The data also documents the dramatic difference between the measured changes in education when using 5 or 10 year time period. The 5 year differentials are entirely dominated by the difference between census and non-census observations.

2.4 Growth regressions: what changes?

2.4.1 OLS estimations

Having analysed the difference in education data according to its source, we will now re-evaluate the GDP regressions. First, we estimate the macro-Mincerian growth equation

as defined by

$$\Delta LGDP_{it} = \gamma_t + \alpha LGDP_{i,t-\tau} + \beta_1 \Delta Edu_{it} + \beta_2 Edu_{i,t-\tau} + \varepsilon_{it} \tag{2.4}$$

where γ_t are time specific effects, $LGDP_{it}$ stands for log real income per worker in country i in period t, Edu_{it} is the average education level, τ is the time span of the data, and ε_{it} is an i.i.d. error term. All variables in changes are annualised.

Table 2.5: The effect of education on growth - annualised OLS estimations

			5 year data		
Variable	Topel(1999)	K&L(2001)	Edu	PEdu	Edu
ΔEdu	0.041**	0.039**	0.0517**	0.0488**	0.0462**
	(0.014)	(0.014)	(0.0137)	(0.0146)	(0.0150)
LagEdu	0.004**	0.004**	0.0035**	0.0037**	0.0036**
	(0.001)	(0.001)	(0.0009)	(0.0010)	(0.0009)
LagLGDP	-0.007**	-0.006*	-0.0060**	-0.0063**	-0.0062**
	(0.002)	(0.003)	(0.0022)	(0.0023)	(0.0023)
Census variables	-	-	-	-	Yes
R^2	0.218	0.197	0.1315	0.1296	0.1414
Observations	608	607	722	722	722
Countries	111	110	97	97	97
			10 year data		
ΔEdu	0.085**	0.086**	0.0882**	0.0789**	0.0758**
	(0.020)	(0.024)	(0.0213)	(0.0222)	(0.0215)
LagEdu	0.004**	0.004**	0.0039**	0.0041**	0.0040**
	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.0009)
LagLGDP	-0.007**	-0.005†	-0.0073**	-0.0076**	-0.0075**
	(0.002)	(0.003)	(0.0021)	(0.0022)	(0.0022)
Census variables	-	-	-	-	Yes
R^2	0.315	0.284	0.2336	0.2240	0.2472
Observations	290	292	353	353	353
Countries	111	110	97	97	97

Notes: Significance levels: † : 10% * : 5% ** : 1%. Robust standard errors in parentheses. The results under Topel(1999) reproduce part of Table 4 in Topel (1999). The results under K&L(2001) reproduce part of Table 3 in Krueger and Lindahl (2001). In this case the number of countries is the maximum number of countries reported by the authors. All variables in changes were divided by the time span in each data. The dependent variable is annualised first-difference real Log GDP per worker, ΔLGDP. All regressions include time effects. In the last column we re-estimate the model using Edu and include the census variables as regressors.

The first two columns of Table 2.5 reproduce estimations from Topel (1999, Table 4), and Krueger and Lindahl (2001, Table 3) [K&L(2001)]. The remaining results are our estimations of equation (2.4) using the two measures of education, Edu and $PEdu$, at different time spans of the data, 5 and 10 years, respectively. The last column of Table 2.5 reproduces the estimations using Barro and Lee's data, and including as regressors the census variables described above. The estimation procedure is OLS, and we report standard errors robust to heteroskedasticity and error correlation within countries.

Similarly to Topel (1999) and Krueger and Lindahl (2001), we also conclude that contemporaneous changes in education have a positive and statistically significant effect on economic growth, which contradicts the findings of Benhabib and Spiegel (1994) and Barro and Sala-i-Martin (1999). For the five year data, our results indicate that the short-run return do changes in education is around 5%, while Topel (1999) and Krueger and Lindahl (2001) results points approximately to 4%. We also conclude that the returns to changes in schooling increases with the time span of the data. Krueger and Lindahl (2001, p.1119) suggest that "the finding that the time span matters so much for the change in education suggests that measurement error in schooling influence these estimates". Our interpretation is that it is the measurement error introduced by the estimation procedure implemented by Barro and Lee that leads to this variation, not the measurement error inherent to the census observations. Approximately 42% of the observations on education in the 10 year data are obtained from census or surveys, while in the 5 year data this figure is only 32%. This difference in data quality is associated with a smaller spurious variation in education in the 10 year data, which may explain why the coefficient on ΔEdu increases with the time span.

The most remarkable feature is that the return to education is lower for our corrected data than for the original data, which is shown to be systematically biased. The standard attenuation bias argument tells that measurement error in an explanatory variable reduces its coefficient, *quod non* in this case. The coefficients for our corrected variable are lower instead of higher, for the 5 year estimation, but in particular for the 10 year estimation. A second thought reveals the reason for this phenomenon. The measurement error reduces the mean of the change in education by some 0.15 year every 5 year period (see Table 2.4). Although the estimated coefficient for the corrected variable is some 6% lower,[8] the estimated effect of education on GDP is 0.6 percentage points larger.[9] So, the effect of the bias on the coefficient is a balance between two forces: introducing the spurious component in ΔEdu reduces the coefficient, while understatement of the average level of ΔEdu pushes

[8] For the 5 year interval, we have $(0.0488 - 0.0517)/0.0517 \simeq -6\%$.

[9] Using $PEdu$ and Table 2.4 one can infer that the return to the average change in education on a 5 year interval is 2.4% ($\simeq 0.0488 * 0.501$), while from Edu we get 1.8% ($\simeq 0.0517 * 0.35$).

up the coefficient. For the 5 year time frame, both forces almost cancel. For the 10 year time frame, the first component is less important (since many census observations are located at the beginning of a decade), so the latter force clearly dominates. When we use Barro and Lee's 10 year original data, returns to changes in education are 8.8%, and very similar to the two comparison studies. However, using our corrected value for education the estimated return is only 7.9%. The systematic measurement error on education identified in the previous section could lead to the overestimation of its coefficient in a growth regression, which is clearly corroborated by the 10 year results. While for Topel, and Krueger and Lindahl, the coefficient more than doubles with the doubling of the time span, the change in our coefficient is smaller, which facilitates the reconciliation between the results for different time spans.

The bias introduced by the perpetual inventory method has a specific structure as described by equation (2.1). This provides an alternative way to account for it in growth regressions. Just for the sake of simplicity, suppose our model is defined as $y_t = \beta_1 + \beta_2 x_t^* + \xi_t$. Assume also that the observed value of x_t^* is $x_t = x_t^* + \phi\omega_t$, where ω_t is observed and represents the lag since last census, and ϕ is a negative coefficient underlying the computation of x_t. The estimated model is $y_t = \beta_1 + \beta_2 x_t + \xi_t - \beta_2\phi\omega_t$. By regressing y_t on x_t the error term is defined as $\xi_t - \beta_2\phi\omega_t$, leading to an omitted variable bias. A simple solution for this bias is to introduce the census information in our income regression. The last column of Table 2.5 reports the results. The growth regression is estimated using Barro and Lee data jointly with the census variables $Before$, $Last$, $LastC$, and $After$. The coefficients on ΔEdu and $LagEdu$ are now slightly smaller than the ones we obtained with $PEdu$. For the 5 year interval data the joint significance test on the census variables yields an $F - statistic$ of 2.81, with a $p - value$ of 0.03, while for the 10 year data the $F - statistic$ is 1.88, with a $p - value$ of 0.12.[10]

A second result, which is identical among the different studies and time spans, indicates that the initial level of education is relevant for economic performance. While the result on ΔEdu supports the human capital interpretation of the role of education in economic growth, this empirical evidence gives also support to the externalities interpretation of the returns to education. Based on our corrected data, $PEdu$, the long-run return to education is $54 - 59\%$.[11] Although this return seems (too) large, we should keep in mind that the effect takes a long time to materialize. The return is at 50% of its long-run value

[10] A further factor that yields overestimation of the effect of education based on the Barro and Lee data is that the variable $Before$ turns out to be a predictor of future growth. The most likely explanation is that holding a census is not an exogenous variable. So countries that initially do not have a census, and later on have, are countries that are likely to have grown faster than average.

[11] That is, 0.0041/0.0076 or 0.0037/0.0063, respectively.

after $75-99$ years. The immediate return is 4.2% for the 5 year time period and 6.5% for the 10 year period.[12] The numbers for the 5 and 10 year time interval are very similar. This puts into question Krueger and Lindahl's interpretation of this difference as being due to an increase in the signal to noise ratio when lengthening the observation period. Lengthening the observation period makes the short return look much like the long-run return, which happens to be substantially higher than the short-run return. In Figure 2.3 the return to education over the first 110 years is depicted. The time path of the cumulated returns to education is very similar for the two time spans, and for the two education variables.

The results indicate that the GDP half–life adjustment ranges between 91 and 110 years,[13] which stresses the idea that whichever externalities are associated with permanent

[12] The immediate return and the half–life can be calculated by assuming that innovations in the education variable are uniformly distributed over the observation period. We do the calculations for $PEdu$, and for the 5 year observation period. First, we calculate the raw estimate of half–life

$$\frac{\ln(2)}{0.0063} = 110.0234$$

Second, we correct for the fact that part of effect is realized immediately. Since the short–run return, S, can be defined as

$$S = L\left(1 - e^{-\lambda t}\right),$$

where L is the long–run return, and λ is the convergence rate to equilibrium, our results imply that

$$0.0488 = \frac{0.0037}{0.0063}\left(1 - e^{-0.0063t}\right).$$

So, the time needed to reach the immediate effect is

$$\frac{\ln\left(1 - 0.0488\frac{0.0063}{0.0037}\right)}{-0.0063} = 13.7695$$

Finally, we take into account the fact that the immediate effect is measured imperfectly, by using a five year time interval. Assuming that the innovation is distributed uniformly, we have to add half of the length of the time interval. The estimated half-life is given by

$$110.0234 - 13.7695 + 2.5 = 98.7539$$

The immediate effect has also to be corrected for the length of the observation period (the longer the observation period, the more the estimated immediate effect will look like the long-run effect). This can be done by taking the time to reach the immediate effect corrected for half the time interval, and using a first order Taylor expansion of the function $1 - e^{-\lambda t}$, λt,

$$0.0063 * (13.7695 - 2.5) = 0.071.$$

Hence, 7.1% of the long-run effect is realized immediately:

$$0.071 \times \frac{0.0037}{0.0063} = 4.2\%.$$

[13] That is, $\ln(2)/0.0076 \simeq 91$ and $\ln(2)/0.0063 \simeq 110$, respectively.

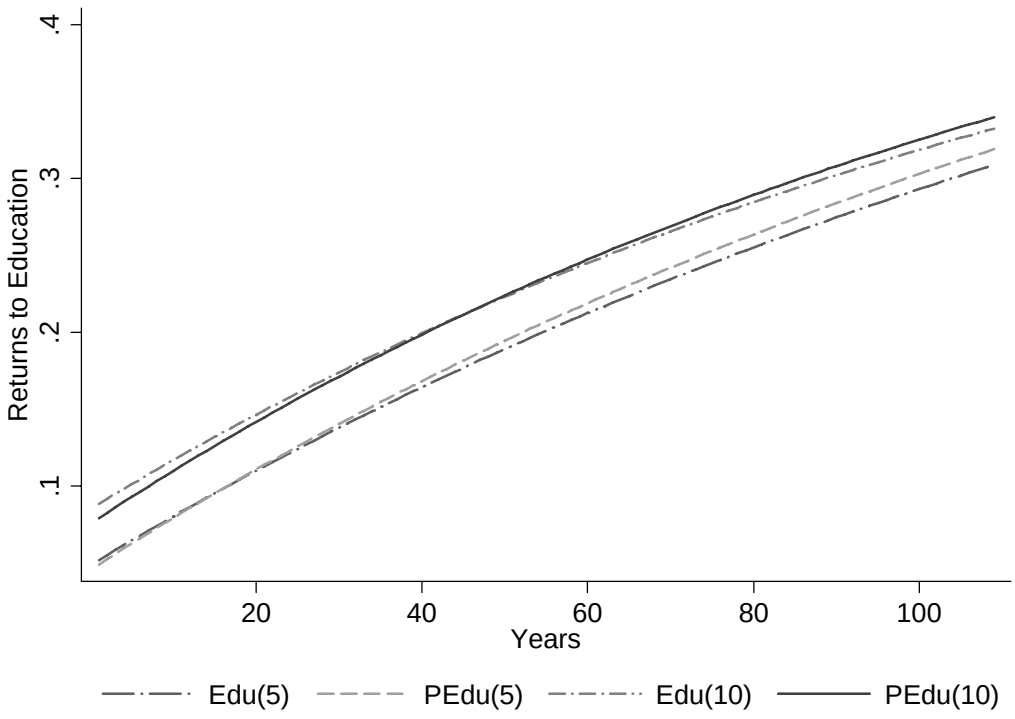

Figure 2.3: Returns to education for the different time spans and education variables

changes in education, they will take a long time before benefiting a given country. The results indicate that the time length of the data sets currently available is too short to identify in a precise way the long-run returns to the investment in education.

Finally, in Table 2.6 we replicate the estimations using Cohen and Soto's (2001) education variable. The differences with Barro and Lee's data occur essentially on the coefficient on changes in education. In this case, the contemporaneous returns to education are around 11%, more than 4 percentage points above our corrected estimates.[14] In the long run, using $EduCS$ indicates a return to education of about 49%, more than 8 percentage points below the values we obtain using Barro and Lee's data.

[14]In our analysis, we are missing 18 countries in Cohen and Soto's data, which are in Barro and Lee sample. The countries are Barbados, Botswana, Congo, Gambia, Guinea-Bissau, Hong Kong, Iceland, Israel, Lesotho, Pakistan, Papua New Guinea, Poland, Rwanda, Slovakia, Slovenia, Sri Lanka, Taiwan, and Togo.

Table 2.6: Growth regressions - comparison with Cohen and Soto data

	10 year data			
Variable	Edu	PEdu	Edu	EduCS
ΔEdu	.0838**	.0664*	.0664*	.1107**
	(.0255)	(.0285)	(.0271)	(.0342)
LagEdu	.0036**	.0039**	.0037**	.0032**
	(.0010)	(.0010)	(.0010)	(.0009)
LagLGDP	-.0063**	-.0066**	-.0065**	-.0065*
	(.0024)	(.0025)	(.0025)	(.0026)
Census variables	-	-	Yes	-
R^2	.2245	.2155	.2474	.2204

Notes: Significance levels: $*$: 5% $**$: 1%. Robust standard errors in parentheses. The dependent variable is annualised first-difference real Log GDP per worker, ΔLGDP. ΔEdu stands for annualised changes in education. All regressions include time effects. The sample includes 300 observations and 79 countries.

2.4.2 Sensitivity analysis: the dynamic model

Adjusting the annualised equation (2.4), we can define the following dynamic model

$$LGDP_{it} = \tau\gamma_t + (1+\tau\alpha)\, LGDP_{i,t-\tau} + \beta_1 \Delta Edu_{it} + \tau\beta_2 Edu_{i,t-\tau} + \tau\eta_i + \tau\varepsilon_{it}, \qquad (2.5)$$

where η_i is the country's specific effect.

In the presence of country's specific effect in equation (2.5), its estimation by OLS and by the usual panel models, fixed or random effects, is inconsistent. The reason is that, by definition, $LGDP_{i,t-\tau}$ in equation (2.5) is always correlated with η_i. One possible solution to overcome this problem is to take first differences in equation (2.5) to eliminate η_i. Arellano and Bond (1991) first-differenced generalized method of moments (GMM) is one of the most applied solutions. Using their procedure avoids the bias introduced by omitted time-invariant variables. However, this solution has poor finite sample properties on bias and precision when "the lagged levels of the series are only weakly correlated with subsequent first-differences, so that the instruments available for the first-differenced equations are weak" (Bond et al., 2001, p.6). Blundell and Bond (1998) show that, in this case, the solution of Arellano and Bond (1991) has a large downward finite-sample bias. This problem occurs when the time series are persistent and the number of time series observations is small. An alternative solution would be to implement a system GMM estimation, for first-differences and levels, as argued by Blundell and Bond (1998). Bond

et al. (2001) argue that this is the best solution to estimate growth regressions.

Table 2.7: Dynamic income regressions – 10 year data

Model	Sys	FD	No-FE	Sys	FD	No-FE
Variable		Edu			PEdu	
LagLGDP	0.780**	0.742**	0.947**	0.915**	0.686**	0.941**
	(0.116)	(0.184)	(0.026)	(0.118)	(0.161)	(0.028)
ΔEdu	0.132**	-0.061	0.073**	0.097*	-0.125	0.073**
	(0.036)	(0.127)	(0.020)	(0.038)	(0.125)	(0.023)
LagEdu	0.133**	-0.071	0.038**	0.094†	-0.134	0.039**
	(0.041)	(0.131)	(0.011)	(0.051)	(0.129)	(0.012)
Wald joint	637.48**	42.59**	6660.31**	658.87**	39.82**	6179.05**
Wald time	104.004**	17.494**	116.953**	85.147**	17.324**	104.197**
Sargan	23.822	11.622	23.623	28.436†	13.541	25.232
Sargan-df	19	12	21	19	12	21
DifSargan	12.200†		12.001	14.895*		11.692
AR(1)	-3.166**	-2.550*	-3.778**	-3.363**	-2.145*	-3.704**
AR(2)	1.557	0.825	1.482	1.243	0.857	1.306
Nobs	350	256	350	350	256	350
Ncountries	94	94	94	94	94	94

Notes: Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. Robust standard errors in parentheses. The dependent variable is real Log GDP per worker, LGDP. All regressions include time effects. Wald is the Wald test on joint significance of the refered set of variables; joint stands for all regressors except dummies; time stands for the time dummies. Sargan and Sargan-df are the Sargan test statistic and degrees of freedom for the test for overidentifying restrictions; DifSargan is the statistic for the test of the validity of the additional set of instruments, when compared with the first-differenced (FD) estimation. Sys stands for system GMM estimation, while No-FE is the system GMM estimation where we assume the absence of a fixed country effect. AR() shows the test for first- and second-order serial correlation in the first-differenced residuals, respectively. The estimation details, including the instrumens used, are described in Section 2.4.2.

Table 2.7, and Table 2.8 in Appendix 2.A, present the results of the estimation of equation (2.5) using the data for 10 and 5 year intervals, respectively.[15] For each education variable, Edu and $PEdu$, and for each time span, we estimate equation (2.5) using the system procedure (Sys), the first-differenced procedure (FD), and a system procedure which assumes the absence of a fixed country effect (No-FE). The dependent variable is real log income per worker ($LGDP$), all regressions include time dummies, and education is treated as a predetermined variable.[16] The instruments for the first-difference equations are the level of $LGDP$ lagged two periods and earlier, and levels of education lagged one

[15]The results in Tables 2.7 and 2.8 are directly comparable with the results in Table 2.5 once we control for the time span of the data. The transformations of the dynamic estimates follow from equation (2.5).

[16]Modelling education as endogenous does not change the main results of this section.

period and earlier. For both variables we use at most 5 lags, following Bowsher (2002)'s suggestion. For the level equations we use first-difference of $LGDP$ lagged one period, and contemporaneous first-difference of education. The estimation of No-FE is similar to system estimation, but in this case the instruments for the equations in levels are not in first-differences but in levels. The reported results are for the 2-step GMM estimation procedure, following the correction proposed by Windmeijer (2005).[17]

We test for the presence of the specific effect following the procedure described in Arellano (2003, p.124). The statistic of the test is the difference in the Sargan test associated with the estimations FD and No-FE, which follows a chi-squared distribution with the number of degrees of freedom given by the difference in the number of instruments in the referred two estimations. We are testing the validity of the additional set of instruments, when compared with the FD estimation. Our results indicate that we do not reject the null hypothesis; i.e., we do not reject the hypothesis that there is no specific country effect.[18] This implies that the results we are lead to interpret are the ones in Table 2.5. OLS estimates are consistent in the absence of unobserved heterogeneity, and they are more efficient. Using equation (2.5) to compare the estimates, we observe that our results for the NO-FE model using $PEdu$ and 10 year data are very similar to the corresponding results reproduced in Table 2.5. The returns to contemporaneous changes in education are 7.3%, while the coefficient on lag education is 0.004, and the coefficient on lag income is 0.006. The comparable figures from the OLS estimation are 7.9%, 0.004 and 0.008, respectively.

Although the results for the estimation of the model Sys are very different when we use Edu and $PEdu$, they become identical when we estimate the model NO-FE. Using $PEdu$ matters when we compare the Sys and the NO-FE estimation, since the results are more similar. In the 5 year data (see Table 2.8 in Appendix 2.A), the results for the system estimation are unreliable with a coefficient on lag income above one. Again, the results for the estimation of NO-FE are similar between the two education variables, with the exception of the coefficient on ΔEdu. Using Edu indicates that the impact multiplier of one year change in education is 4.7%, while using $PEdu$ the equivalent value is 3.7%. As before, the results for the 10 year data seem to be more stable.

[17] We used the Ox version of DPD (Doornik et al., 2002) to obtain the results in Tables 2.7 and 2.8.

[18] The No-FE type of regression has two more instruments when compared with the system estimation. The reason is that in the first case we use an extra period in the level equations.

2.5 Final remarks

Our analysis of Barro and Lee (2001) education data reveals that there is a systematic difference between the information collected from census or surveys, and the education data that results from the perpetual inventory method. On average, this method underestimates education by about one fifth of a year every five year period. This has an impact on the results for the growth regressions. Once we control for the source of information, and we take into account measurement error, we conclude that both the level and the change in education are relevant for the growth process. However, alternative specifications and data intervals make a difference for the size of the effects. Further research is needed in order to make proper use of the knowledge on the systematic difference between census and non-census data.

Following Teulings and van Rens (2006), it would be important to take into account second order effects on education. The re-estimation of the data on education is another alternative for future work. Using both the backward and the forward flow, the missing values can be reestimated using the average of both predictions, not only the weighted average between the linear interpolation and the forward prediction. However, the estimation of the missing values after the last census, and before the first census, would still be estimated the same way. It would also be important to estimate educational values taking into account different survival rates according to the educational attainment. On this topic, Barro and Lee (2001, p.545) state that "the limitation of the data on age-specific education levels and mortality rates by age group do not allow us to compute specific mortality rates of population by levels of education."

Appendix to Chapter 2

2.A Additional results and census example

Table 2.8: Dynamic income regressions – 5 year data

Model	Sys	FD	No-FE	Sys	FD	No-FE
Variable		Edu			PEdu	
LagLGDP	1.038**	0.909**	0.973**	1.043**	0.899**	0.973**
	(0.054)	(0.083)	(0.011)	(0.049)	(0.073)	(0.011)
ΔEdu	0.037†	0.028	0.047**	0.034	0.013	0.037*
	(0.020)	(0.028)	(0.016)	(0.021)	(0.034)	(0.016)
LagEdu	0.015	0.017	0.017**	0.014	0.008	0.016**
	(0.018)	(0.028)	(0.005)	(0.017)	(0.033)	(0.005)
Wald joint	1823.20**	159.63**	25416.69**	1862.55**	173.93**	26313.78**
Wald time	103.707**	25.615**	89.606**	68.701**	24.234**	81.924**
Sargan	76.094	49.283	73.249	73.300	54.908	73.794
Sargan-df	63	48	65	63	48	65
DifSargan	26.811*		23.966	18.392		18.885
AR(1)	-3.576**	-4.144**	-3.611**	-3.572**	-3.968**	-3.602**
AR(2)	1.080	1.068	1.089	1.051	1.021	1.055
Nobs	722	625	722	722	625	722
Ncountries	97	97	97	97	97	97

Notes: Significance levels: † : 10% ∗ : 5% ∗∗ : 1%. Robust standard errors in parentheses. The dependent variable is real Log GDP per worker, LGDP. All regressions include time effects. The estimation details, including the instrumens used, are described in Section 2.4.2. See note to Table 2.7 for details on the reported tests.

Table 2.9: Example of census variables for Figure 2.1

Year	*Census*	*Before*	*Last*	*LastC*	*After*
1960	0	1	0	0	0
1965	1	0	0	0	0
1970	0	0	1	0	0
1975	0	0	2	0	0
1980	1	0	3	3	0
1985	1	0	1	1	0
1990	0	0	1	0	0
1995	1	0	2	2	0
2000	0	0	0	0	1

Chapter 3

Economic cross-country spillovers of education

3.1 Introduction

Computations based on country level data provided by Heston et al. (2002) and Barro and Lee (2001) show that for the year 2000 the difference in log income per worker between the top and bottom deciles is just above 3, while the difference between top and bottom deciles of average schooling is 7.4 years. Within a very simplified framework, and if we interpret such difference as causal, the social return to education would be some 50%.[1] Private returns to schooling are usually bounded between 6% and 10%, which would imply externalities of 40% to 44%.

Using data for the United States (US) and aggregating education at the state level, Acemoglu and Angrist (2000, p.40) conclude that social returns are bounded between 8% and 9%; they claim that "this is clearly too small to rationalize the steep relationship between average schooling and output per worker". They conclude that external returns to education are not significant, both statistically and economically. The authors, however, do not completely rule out external returns. They state that their "strategy identifies local effects, missing external returns that raise wages nationwide" (Acemoglu and Angrist, 2000, p.48). Also the confidence intervals based on their standard errors would not reject external returns bounded between 1% and 3%.

Externalities of education can assume different forms. According to Krueger and Lindahl (2001), if human capital is expanded at higher levels we could observe spillovers in the form of technological progress, and gains in productivity, while if it is expanded at lower levels we could observe lower crime rates and welfare participation. By using child

[1] Note that $0.5 \simeq \exp(3/7.4) - 1$. Computations based on a sample of 93 countries with information on income and education.

labour and compulsory attendance laws as instruments, Acemoglu and Angrist's (2000) focus in the lower part of the schooling distribution. Looking into the other side of the distribution, Moretti (2004a) studies spillover effects of the share of college graduates at the city level, and concludes in favor of positive and significant externalities.[2] When the analysis is implemented at the state level the direction of the results is unchanged, but the size of the spillovers is smaller. Another difference between Acemoglu and Angrist (2000) and Moretti (2004a) is that the first authors draw their main conclusions from the period 1960-1980, while the second focuses his attention in late 1970s and early 1990s. Moretti (2004b) also finds significant human capital externalities by focusing on the productivity of manufacturing establishments.

If we aggregate even more the level of analysis, and look into the growth literature, the conclusion would be that the role of human capital in the growth process is not clearly identified. As we saw in Chapter 2, the returns to education can vary significantly across studies, relying very much on the quality of the data used. As stated by Temple (2001, p.905), "the empirical evidence that education matters for growth is surprisingly mixed."

From a theoretical perspective, Glaeser et al. (2003) argue that the existence of spillovers creates a social multiplier, which implies that aggregate coefficients are greater than individual ones. This multiplier rises with the level of aggregation. In the presence of social interactions the use of aggregate data to make inferences about individual behavior is not valid, since "aggregate relationships will overstate individual elasticities. [However, the authors state that for many purposes...] researchers actually want the aggregate coefficient that includes both the individual level response and the social multiplier" (Glaeser et al., 2003, p.345-346). Sianesi and van Reenen (2003, p.166) put it clearly by saying that "one justification for the macro growth regressions is precisely their potential ability to capture economy-wide indirect or spill-over effects from educational investments."

By focusing on countries, and dealing explicitly with neighbouring education, we aim at identifying and quantifying global economic externalities of education. Not only we want to identify social interactions between individuals organized within the same country, but also we want to identify interactions across countries. We look for returns to education that could justify observed investments in education. Similarly to Teulings and van Rens (2006), we argue in our analysis that there are diminishing returns to education. If this is the case, then different countries would show different returns to education. Sianesi and

[2] According to Moretti (2004a), a one percentage point increase in the proportion of college graduates would raise wages between 0.6% and 1.2%. To translate these results into returns to years of education, we need to assume that "the share of college graduate[s] inceases by 25 percentage points and the share of high-school graduates decreases by the same amount" (Moretti, 2004a, p.195). A direct computation points to external returns of 15% to 30%. We have to account for the fact that actual change in the share of college graduates is below one percentage point.

van Reenen (2003) observe that schooling returns are generally higher in less developed countries than in the OECD. Our interpretation is that this result can be explained by diminishing returns to education and high levels of integration between OECD economies.

Another point we stress is that part of the returns to education are only identifiable at the group of countries level, and not at the individual country level. Economic distance, a concept used by Moretti (2004b) and Conley and Ligon (2002), among others, is key to our analysis.[3] Moretti (2004b, p.674) finds that "within a city, the magnitude of the spillovers decline with economic distance." The author argues that industry aggregate human capital benefits more those plants that are economically closer. Extending the arguments for country data, we use bilateral trade to measure the interaction between countries, define neighbourhood, and build the variable neighbouring education. International trade is an important channel for interactions between countries and fosters knowledge diffusion. Foreign research and development investment, as well as the technological diffusion, are also associated with it.

When testing for externalities an important concern is raised by workers' mobility. Acemoglu and Angrist (2000) and Moretti (2004a,b) argue that it causes an identification problem. If mobility across cities/states is justified by cities/states' unobserved factors, estimates of external returns would be biased. By dealing with a different level of aggregation we reduce this problem. Our argument is that between cities, or states, migration of workers is much easier and significant than across countries, which renders it a smaller problem when aggregation occurs at the country level.

Our results indicate that the US have lower returns to education when compared to other countries, which can be explained by diminishing returns to education. This makes it more difficult to find externalities when confining the analysis to the US. Over a 40 year time interval, our results show, for example, that the US total returns to education are below 17%, compared to 25% for the United Kingdom (UK). It follows that using the US cities, or states, has limitations when we want to test for global externalities, or to understand the cross-country association between average education and average income. By using country level data our results point to considerable overall returns to education. On average, long-run returns to education can be at least as high as 66%, although its half–life period is above 65 years. Within the 40 year interval of our data, social returns are bounded between 6% and 17%. After accounting for private returns, education spillovers are still considerable.

The chapter is organized as follows. In Section 3.2 we define our measure of interaction

[3] One of the economic distance measures used by Moretti (2004a) is based on input-output tables. The author wants to "capture interactions between industries that arise from exchanging goods and services during the production process" (Moretti, 2004b, p.671).

between countries. With these data we build the neighbouring education variable used in the following sections. The data is described in Section 3.3 and the empirical analysis is discussed in Section 3.4. We first estimate the proximity measure (Section 3.4.1), and then we test for the presence of country effects (Section 3.4.2). As we conclude in favor of absence of country effects, we proceed by using least squares procedures. In Section 3.4.3 we test for spillovers, and quantify the different returns to education. Finally, in Section 3.5, we discuss some concluding remarks.

3.2 How to measure the interaction between countries?

3.2.1 The model

Suppose that we can order all workers according to their 'economic' distance to a particular reference worker, and let x be the rank of a particular worker in the ordering. Hence, $x = 0$ is the reference worker himself, $x = 1$ is his closest neighbour. We use the concept 'economic' rather than 'physical' distance to stress that we are interested in economic linkages.[4] Let the log output of a worker living at location 0 and time t, $y_t(0)$, satisfy

$$\begin{aligned} y_t(0) &= \alpha H_t(0) + \beta \int_0^\infty g(x) H_t(x)\, dx, \qquad (3.1) \\ \int_0^\infty g(x)\, dx &= 1, g(x) > 0,\ g'(x) < 0, \end{aligned}$$

where $H_t(x)$ is years of education of a worker at location x and time t, and $g(x)$ is a weighting function. The first term, $\alpha H_t(0)$, is the classical Mincerian private return to education. The second term measures the spillover effect, or externality, of the education of a given worker neighbours on his own output. Since the weights $g(x)$ of all people sum to unity, β measures the total spillover of his neighbours.

The actual data are grouped by country. Let x_j be the distance to the worker in country j who is closest to location $x = 0$. For simplicity, we assume that the locations in other countries are closed and non-overlapping subsets of the domain of x, so that a worker lives in country j if and only if $x \in [x_j, x_{j+1})$.[5] By definition, $x_0 \equiv 0$, so that

[4] For example, if we think in terms of countries, though Tibet is physically closer to Nepal than the United Kingdom to the United States, the latter two might be closer in economic terms. The Himalaya might be literally an unsurpassable barrier of trade, while the common language and cultural tradition, and the cheap transport options offered by the Atlantic, make the US and the UK close neighbours.

[5] This is not an innocuous assumption; e.g., two countries bordering to country 0 have overlapping subsets of the domain of x. We ignore this complication here.

country $j = 0$ is the home country of worker $x = 0$. Hence, we can approximate the final term as a weighted sum of average education levels in countries:

$$\begin{aligned} \int_0^\infty g(x) H_{xt} dx &\cong \sum_{j=0}^{I-1} g_{0j} \overline{H}_{jt}, \\ g_{0j} &\equiv \int_{x_j}^{x_{j+1}} g(x) dx \cong g(x_j)(x_{j+1} - x_j), \end{aligned} \tag{3.2}$$

where $\overline{H}_{jt}$ is the average education level of the labour force in country j and where I is the number of countries. The approximation in the second line shows that the weight g_{0j} captures two aspects of country j: its distance, $g(x_j)$, and its size, $(x_{j+1} - x_j)$. The closer by another country, or the larger that other country, the greater the spillover effects. Furthermore, note that the home country of worker $x = 0$ is also included in the summation. This measures the spillover effect of a worker compatriots on his own productivity. The more compatriots a worker has, the larger their joint impact on his output. Hence, the larger the home country, the larger its role in the summation, and hence the smaller the role of other countries. Accounting for within country externalities is crucial, because the size of countries differs widely. We cannot expect to find much externalities for the US, since the size of that economy is such that we expect most of the externalities to be exploited within the country.

Replacing equation (3.2) in (3.1), averaging over the home country,[6] and, for the sake of generality, replacing the index 0 by the index i, yields:

$$\overline{y}_{it} = \delta + \alpha \overline{H}_{it} + \beta \sum_{j=0}^{I-1} g_{ij} \overline{H}_{jt}, \tag{3.3}$$

where $\overline{y}_{it}$ is mean log output per worker, or log GDP per worker. Since $\int_0^\infty g(x) dx = 1$, then $\Sigma_j g_{ij} \cong 1$. So, if the education level of the whole world, except of a given worker, increases by one year, his output goes up by 100 β%.

3.2.2 Measuring proximity

How can the weights g_{ij} be determined? Our proposal is to use bilateral trade flows as a proxy for the link between countries. The volume of bilateral trade captures two features of two countries, the size of both countries and their proximity; i.e., the larger

[6] Formally, averaging over the home country is problematic, since rank ordering x differs across the workers in the home country; e.g., workers living just before the border to country $j = 1$ have a different ordering than those at the centre of the country. Again, we ignore this complication.

the trading partner, the larger the bilateral trade with that partner, and, almost by definition, the closer by the trading partner is in economic sense, a higher bilateral trade is expected. Both effects matter for estimating externalities. For instance, we cannot expect Luxemburg to have large external effect on the Dutch economy; though it is close by, its limited size does not allow for large externalities.

Consider the following economic model for bilateral trade:

$$\bar{t}_{ij} = \gamma_0 + \gamma_1 \Delta_{ij} + \gamma_2 \Delta_{ij}^2 + f_i + f_j + \xi_{ij}, \tag{3.4}$$

where $\bar{t}_{ij}$ is log trade (the sum of the flows from i to j and from j to i) between countries i and j, averaged over the observation period; f_i is a country-specific effect for country i, capturing both the size of the GDP of country i and its general openness to trade; Δ_{ij} is the physical distance between country i and j; ξ_{ij} is an error term capturing unmeasured differences in proximity between country i and j, due to e.g. geographical, cultural or political factors; and γ_0, γ_1, and γ_2 are unknown parameters. We estimate equation (3.4) by OLS. Next, we filter the effect of the home country's education level out of the fixed effect by running the regression:

$$f_i = \mu + \beta \overleftrightarrow{S}_i + \tau_i, \tag{3.5}$$

where $\overleftrightarrow{S}_i$ denotes the average years of schooling among the workforce $\overline{S}_{it}$, averaged over the observation period, and τ_i is an error term. This way we correct the size of a country for the effect of education to avoid double counting. Then, the proximity between country i and j, w_{ij}, is defined by

$$\begin{aligned} w_{ij} &= \gamma_1 \Delta_{ij} + \gamma_2 \Delta_{ij}^2 + \tau_i + \tau_j + \xi_{ij} \\ w_{ii} &= 2\tau_i \Rightarrow \xi_{ii} \equiv 0. \end{aligned}$$

The weight w_{ij} captures: (i) the physical proximity of country j to country i, $\gamma_1 \Delta_{ij} + \gamma_2 \Delta_{ij}^2$; (ii) non-physical factors affecting the bilateral trade between both countries, like a common language or juridical system, ξ_{ij}; and (iii) the size and general openness of both countries i and j, measured by the filtered fixed effects τ_i and τ_j. To the extent that Δ_{ij} and Δ_{ij}^2 are orthogonal to the other regressors, we could have just omitted this variable in equation (3.4), since its effect would have been captured by ξ_{ij}, and it would have shown up in the weights w_{ij} anyway. However, capturing the impact of physical distance is both instructive and it constitutes a more robust way of establishing w_{ij}. Since the distance to the home country, Δ_{ii}, is zero by construction, and ξ_{ii} is zero by definition,

the diagonal elements of the weighting matrix, w_{ii}, measure only the effect of the size and general openness of country i.

Each weight w_{ij} can be either positive or negative, whereas the weights g_{ij} can only be positive. Hence, the actual weights that we use are an exponential transformation of w_{ij}:

$$g_{ij} = \frac{\exp(\lambda w_{ij})}{\Sigma_j \exp(\lambda w_{ij})}, \tag{3.6}$$

where the parameter λ will be estimated. The parameter λ measures the elasticity of externalities with respect to bilateral trade: if λ is equal to unity, a 1 % smaller bilateral trade reduces the size of the externality by 1 %.[7] The division by $\Sigma_j \exp(\lambda w_{ij})$ makes that $\Sigma_j g_{ij} = 1$. Note that this definition implies that g_{ij} is invariant to a mean shift in the weights w_{ij}, as it should be (e.g., the currency in which trade is measured or the rate of inflation of that currency may not affect our conclusions). Unlike the w_{ij}'s, the g_{ij}'s are not symmetric: $w_{ij} = w_{ji}, g_{ij} \neq g_{ji}$. This fits, since the size of externalities per capita is not symmetric: for instance, Germany has a larger impact on Luxemburg than Luxemburg on Germany, simply because Luxemburg is small and Germany is big. Nevertheless, the normalization $\Sigma_j g_{ij} = 1$ might not be the appropriate thing to do. Suppose that an economy, let us say Eastern Island, is not well attached to the rest of the world economy. We measure that by the fact that w_{ij} is small, not only because Eastern Island itself is a small economy, as measured by τ_i being small, but also because it is too far away from other economies, as measured by Δ_{ij} and ξ_{ij}, to benefit from their externalities. From this perspective, an appropriate definition of g_{ij} would be:

$$g^*_{ij} = \exp(\lambda w_{ij}), \tag{3.7}$$

which does not normalize the sum of the weights to unity. Now, the term $\Sigma_j g^*_{ij} \overline{S}_j$ measures two things simultaneously: first, the connection of country i to the rest of the world, $\Sigma_j g^*_{ij}$; and second, whether the 'economic neighbours' of i (the countries j with large g^*_{ij}) are better educated than average. This is as it should be: we can only benefit from education spillovers if we have many neighbours which are moreover well educated. Note that this definition implies that g^*_{ij} is invariant to a mean shift in the τ_i's. We return to the distinction between g_{ij} and g^*_{ij} below, when discussing the empirical results.

[7] From equation (3.6) we obtain $\frac{dg_{ij}}{dw_{ij}} = \frac{\lambda e^{\lambda w_{ij}}\left[\Sigma_j e^{\lambda w_{ij}} - e^{\lambda w_{ij}}\right]}{\left[\Sigma_j e^{\lambda w_{ij}}\right]^2} = \lambda g_{ij}(1 - g_{ij})$, or $\frac{dg_{ij}}{g_{ij}} \frac{1}{dw_{ij}(1-g_{ij})} = \lambda$.

3.2.3 The regression model

Equation (3.3) allows only for direct spillover effects of education in country j on GDP of country i. However, there might be indirect spillovers, if education in country j raises the output of country j, and subsequently the extra GDP in country j has spillover effects on country i. In the current analysis, we are not concerned on whether the education in country j affects GDP in country i directly or indirectly. In both cases, it implies that education of one person has spillover effects on the productivity of others. Hence, we can write the income equation as:

$$\overrightarrow{y}_t = \overrightarrow{\delta} + \alpha \overrightarrow{S}_t + B \overrightarrow{S}_t + \overrightarrow{\varepsilon}_t, \tag{3.8}$$

where $\overrightarrow{y}_t$, $\overrightarrow{\delta}$, and $\overrightarrow{S}_t$ are vectors with elements for each country i. The matrix B can be written as $\beta G(\lambda)$, where $G(\lambda)$ is a $I \times I$ matrix with elements g_{ij} given by either equation (3.6) or (3.7). The total number of parameters to be estimated is I (for $\overrightarrow{\delta}$)+3 (α, β, and λ).

Teulings and Van Rens (2006) find clear evidence of diminishing returns to education, by entering second-order effects of average years of education. The second-order term comes in negatively. Furthermore, they find strong evidence for dynamics in the return to education, with the long-run return being much higher than the short-run return. Adding both features and time to the model yields:

$$\begin{aligned} \overrightarrow{y}_t &= \overrightarrow{\delta} + \rho \overrightarrow{y}_{t-1} + \alpha_0 \overrightarrow{S}_t + \alpha_1 \overrightarrow{S}_{t-1} + \alpha_{02} \overrightarrow{S^2}_t + \alpha_{12} \overrightarrow{S^2}_{t-1} \\ &\quad + \beta_0 G(\lambda) \overrightarrow{S}_t + \beta_1 G(\lambda) \overrightarrow{S}_{t-1} + \beta_{02} G(\lambda) \overrightarrow{S^2}_t + \beta_{12} G(\lambda) \overrightarrow{S^2}_{t-1} + \gamma_t \overrightarrow{\iota} + \overrightarrow{\varepsilon}_t, \end{aligned} \tag{3.9}$$

where $\overrightarrow{S^2}_t$ is a vector with squared average years of education as its elements; and $\overrightarrow{\iota}$ is a vector of ones. The causal interpretation of the Mincerian terms $\alpha \overrightarrow{S}$ is open to discussion: does education cause GDP, or is it the other way around? The more important question for this chapter is, however, whether reverse causality is also a problem for the spillover terms $\beta G(\lambda) \overrightarrow{S}$. We claim this problem is much smaller. There are two potential problems here: (i) the endogeneity of $\overrightarrow{S}$, and (ii) the endogeneity of the trade flows that determine $G(\lambda)$.[8] Regarding the first issue, reverse causality is much less of a problem for the spillover terms than for the Mincerian terms. In fact, it is quite difficult to think of a mechanism by which a change in y_{it} affects S_{jt} for country i neighbour countries j, unless there is a spillover effect in the first place, by which the change in y_{it} affects y_{jt} for the neigbour countries. However, the effect of this change on S_{it} will be much stronger

[8] See equation (3.4).

than on S_{jt}, making it unlikely that the spillover term measures reverse causality, in particular, when we simultaneously control for the Mincerian effects. The second issue is the potential endogeneity of $G(\lambda)$: a change in y_{it} is quite likely to affect trade flows, and hence $G(\lambda)$. However, this is not a problem, since we do not allow for time variation in the weighting pattern $G(\lambda)$. The spillover terms are only identified by their variation over time. Without time variation, the spillover terms would be swallowed by the fixed effects, $\overrightarrow{\delta}$. Since time variation identifies the terms, and $G(\lambda)$ is invariant over time, reverse causality cannot explain the significance of the spillover terms.

It is interesting to speculate what set of parameters would fit most easily a credible story on knowledge spillovers of education. One would expect that spillover effects or externalities are measured by the $\beta G(\lambda)\overrightarrow{S}$ terms, since there is no reason why spillovers would be confined to the home country. Such limitation to the home country might apply to large countries, where all spillovers are exploited within the country's borderlines, but this is unlikely to apply to a small country like the Netherlands, unless the rate of decay of spillovers, λ, is extremely high. Since our weights g_{ij} capture this difference between small and large countries, the $\beta G(\lambda)\overrightarrow{S}$ terms are most appropriate to capture spillovers. Since knowledge spillovers affect the growth of productivity more than its level, one would expect these terms to have effect predominantly in the long run. Given that spillovers are captured by the $\beta G(\lambda)\overrightarrow{S}$ terms, one would expect the $\alpha\overrightarrow{S}$ terms to capture the private gains to education. Hence, the long- and short-run returns should be about equal (as it measures the effect on the own level of productivity). The marginal return to education must be about equal to the Mincerian rate of return to be credible as an internal return; otherwise it is hard to understand why people would not invest in more education, in particular in countries with well established capital markets and government institutions to facilitate the investment in human capital. There might be declining marginal returns to education, as measured by a negative second order effect, reflecting imperfect substitution on the demand side (see Teulings and Van Rens, 2006). This implies that the equality of marginal return to the Mincerian return must hold for the mean level of education over countries and years in the sample.

3.3 Data

The trade weights are computed based on the data set released by Rose (2004), which contains data on bilateral trade for a sample of 178 countries. Although it spans from 1948 till 1999, we use information for the period 1950–1999 since there are many missing values for the first two years. We have non–zero information on approximately 77% of all

possible pairs among the 178 countries. For pairs with missing information on bilateral trade w_{ij} is set to $-\infty$, such that its exponential is zero.[9] The procedure used to compute the weights takes into account the relevance of trade for each economy. The sorting of countries is based on the gravity model defined by Rose (2004). The variable distance is measured as the great-circle distance between countries in miles, and it is also taken from the data made available by Rose (2004).

Penn World Table 6.1, in Heston et al. (2002), is the source of information for countries' income. Specifically, we use real Log GDP per worker. The education variable is provided by Barro and Lee (2001). We use information on individuals aged 15 years and older. We use average years of education within a country, instead of using enrolment data commonly used in recent studies. Although some efforts have been made to build more complete measures of human capital, namely by Mulligan and Sala-i-Martin (2000), Pritchett (2001), and Woessmann (2003), we concentrate our analysis on the role of education.

In Appendix 3.B, Table 3.10, we provide the list of the 75 countries used in the estimations. Our sample consists of a balanced panel with 5 observations per country, corresponding to a total of 375 observations. The data spans from 1960 till 2000, on 10 year intervals. Among the sample of countries, 16 are African, 3 are from the Middle East, 9 from North America, 10 from South America, 4 are Caribbean, 14 are Asian, 17 are European, and 2 from Oceania. Furthermore, 13 countries are members of the European Union, and 24 are OECD members. Due to data restrictions, none of the members of the former Soviet Union is included in our analysis, and Romania is the only Eastern European country considered. This distribution shows the diversity of countries we have in our sample, and how it represents the different stages of economic performance throughout the world.

Between 1960 and 2000 the highest growth in income per worker occurs in Ireland, Japan, Thailand, South Korea, and Hong Kong, with figures of 161%, 162%, 176%, 211%, and 221%, respectively. The group of countries with the lowest achievement includes Nicaragua, Niger, Venezuela, Mali, and Zambia, all with negative changes in their income per worker over the entire sample period (that is, -53%, -41%, -35%, -31% , and -24%, respectively).

[9] Linders and de Groot (2006) argue that in the presence of zero trade flows the use of a sample selection model should be considered. An alternative solution is to use a sample that excludes the zero flows. In their particular application, and for this alternative solution, the authors conclude that "there is only marginal indication that OLS is biased downwards due to sample selection bias" (Linders and de Groot, 2006, p.9). We have adopted the exclusion of zero trade flows in the current analysis, although further checks should be implemented in order to assure that sample selection does not bias significantly our results.

The variables are described in Table 3.1 for each one of the years, and for the entire sample. On average, income increased 1.8% per year, while education increased 0.07 years. Neighbouring schooling, $G(\lambda)\overrightarrow{S}$, is computed as a weighted average of education in the neighbouring countries in period t, where g_{ij} are the weights, and we assume just for the descriptive statistics that $\lambda = 1$. In a given period,

$$G(1)\overrightarrow{S}_t = \Sigma_j g_{ij} S_{jt} = \Sigma_j \left(\frac{\exp(w_{ij})}{\Sigma_j \exp(w_{ij})} S_{jt} \right) \tag{3.10}$$

Table 3.1: Summary statistics

Variable	1960	1970	1980	1990	2000	Total
$\overrightarrow{y}$	8.856	9.176	9.373	9.444	9.573	9.285
	(0.956)	(1.006)	(1.025)	(1.061)	(1.119)	(1.059)
$\overrightarrow{S}$	3.919	4.474	5.335	6.068	6.677	5.295
	(2.524)	(2.598)	(2.721)	(2.746)	(2.783)	(2.847)
$G(\lambda = 1)\overrightarrow{S}$	3.521	4.096	4.898	5.640	6.340	4.899
	(1.048)	(1.077)	(1.173)	(1.085)	(1.106)	(1.493)
$\Delta\overrightarrow{y}$		0.032	0.020	0.007	0.013	0.018
		(0.021)	(0.023)	(0.024)	(0.021)	(0.024)
$\Delta\overrightarrow{S}$		0.056	0.086	0.073	0.061	0.069
		(0.070)	(0.083)	(0.073)	(0.037)	(0.068)
$\Delta G(\lambda = 1)\overrightarrow{S}$		0.058	0.080	0.074	0.070	0.07
		(0.015)	(0.028)	(0.023)	(0.011)	(0.022)

Notes: The sample contains 10 year data, 75 countries, and 375 observations. This table reports means and standard deviations in parentheses. The last column shows statistics for the entire sample. The symbol Δ indicates annualised changes in the corresponding variable; λ is the decay parameter used in the computation of neighbouring schooling.

Neighbouring education has a lower dispersion than countries own education. From Table 3.1 we conclude that average income per worker has increased more than 70% between 1960 and 2000, while average education has increased 2.8 years. The pattern of neighbouring education is similar to that of countries.

3.4 Empirical analysis

3.4.1 Estimating the proximity measure

Regression coefficients for equations (3.4) and (3.5) are reported in Tables 3.2 and 3.3, respectively. The physical distance, Δ_{ij}, contributes substantially to the explanation of the trade patterns between countries, but the standard deviation of ξ_{ij} shows that remains a substantial role for other factors (Table 3.2). Would the trade vary proportionally to GDP with a change in the average education level, the coefficient on $\overleftrightarrow{S}_i$ in equation (3.5) should have been the same as that in a regression on log GDP. The big coefficient suggests that trade varies substantially with education (Table 3.3). The standard deviation of the country fixed effect, f_i, is higher than that of τ_i, since education explains part of the difference in GDP between countries, and hence the level of their trade. The two countries with the biggest $\hat{\tau}_i$ are Brazil (3.281) and China (3.984), while the two countries with the smallest $\hat{\tau}_i$ are Barbados (-4.600) and Iceland (-3.013).[10] This pattern is as expected: the filtered fixed effect is highly correlated with the size of the country.

Table 3.2: Log average trade regression

Distance	-0.861***
	(0.024)
Distance^2	0.047***
	(0.002)
Constant	21.263***
	(1.714)
Observations	12150
R^2	0.748
Standard deviation: ξ_{ij}	1.688
Standard deviation: f_i	2.165

Notes: Reported coefficients are significant at the 1% level ($***$). Standard errors in parentheses. The dependent variable is log average trade. The regression includes a set of dummies for the country and trade partner, which are jointly significant at the 1% level.

Assuming $\lambda = 1$, Tables 3.11 and 3.12 in Appendix 3.B show weights g_{ij} for some selected trade partners for the United States and the Netherlands. Consider, for example, the US. Its own weight is 49%, while France, Italy, Mexico, Japan, and Brazil are its

[10] The two countries with the highest $\hat{f}_i$ are The United States (21.263) and Russia (20.641). On the other extreme are Bhutan (11.099) and Tonga (11.052).

most important neighbours. If we take The Netherlands, the most important countries are France, Belgium and Italy, and only then the country itself. Again, this fits our expectation. For large countries, most trade happens within its boundaries, so the own country is the main trading partner. Small countries have a limited number of trading partners within the country, and they trade mainly with the rest of the world.

By combining the results obtained in the estimation of equations (3.4) and (3.5) we build the weights defined in Section 3.2.2. These weights are used in the estimation of equation (3.9), which results are discussed in Sections 3.4.2 and 3.4.3.

Table 3.3: Estimated country effects *vs* education

$\overleftrightarrow{S}$	0.547***
	(0.055)
Constant	13.019***
	(0.309)
Observations	110
R^2	0.479
Standard deviation: τ_i	1.526

Notes: Reported coefficients are significant at the 1% level ($***$). Standard errors in parentheses. The dependent variable is the estimated country effect obtained from the estimation of equation(3.4).

3.4.2 Testing for the presence of country effects

The presence of the fixed effect $\overrightarrow{\delta}$ in equation (3.9) implies that it cannot be estimated consistently by GLS procedures, given the correlation with the lagged endogenous variable. In this section we discuss the presence of such fixed effects. Endogeneity concerns relating to our weights might also arise given the potential endogeneity of trade in a growth model, as discussed in Section 3.2.3. However, the validity of this critic is minimized because we do not use bilateral trade as the weight. To implement the different tests we have imposed the restriction $\lambda = 1$, and estimate four flavors of equation (3.9). First, we impose $\beta_0 = \beta_1 = \beta_{02} = \beta_{12} = 0$, specification I. It follows specification II, where we estimate all parameters. In the next two sets of estimates we restrict $\beta_{02} = \beta_{12} = 0$, specification III, thus eliminating nonlinear terms in neighbouring education, and in a final stage the restriction becomes $\beta_0 = \beta_{02} = \beta_{12} = 0$, specification IV, further excluding contemporaneous neighbouring education.

Along side with OLS, we use GMM procedures outlined in Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998) to estimate equation (3.9). For the GMM type of regressions we treat own education, $\overrightarrow{S}$, and neighbours' education, $G(\lambda)\overrightarrow{S}$, as predetermined variables, such that both are independent of $\overrightarrow{\varepsilon}_s$ for $s \geq t$. The instruments used in each regression are described in the discussion that follows. For the GMM system procedures we report the results for the 2–step GMM estimation, which includes the corrected covariance matrix proposed by Windmeijer (2005). Finally, we implement the bias corrected least squares dummy variable estimator (LSDVC) discussed in Bun and Kiviet (2003), and further extended by Bruno (2005a).

A preliminary test is the following. If there are country effects we should observe serial correlation in the error term. To test for it, we first estimate the model defined in equation (3.9) by OLS, and retrieve the estimated residuals. Second, we estimate a regression of these residuals on the lagged residual and all explanatory variables used in the first regression.[11] When we exclude neighbouring education from our model, specification I, the coefficient on lag residual is 0.155, with a standard error of 0.082. This is statistically significant at the 10% level. The remaining explanatory variables included in the regression are jointly not significant, with an F statistic of 0.031. Given the small number of observations per country we also compute the bootstrap $p - value$ for the coefficient of the lag error.[12] The result is 0.221, and the null hypothesis of absence of country effects is not rejected. When we estimate specification II the coefficient of the lagged residual is 0.144, with a standard error of 0.085. All remaining explanatory variables are jointly not significant, with an F statistic of .026. For this estimation, the bootstrap $p - value$ is 0.234. Once again this preliminary tests does not reject the null hypothesis that there are no country effects. Using specifications III and IV similar results are achieved.[13]

Tables 3.4, 3.5, 3.6, and 3.7, report further checks on country effects. The four sets of results are all based on the same empirical strategy, but considering different sets of regressors. In the first column of all tables we estimate the model by OLS. It follows the dynamic estimations, namely Arellano and Bond (1991) and Blundell and Bond (1998), in columns (2) and (3). We have been parsimonious in the use of instruments, as suggested by Bowsher (2002). In columns (2) we use $\overrightarrow{y}_{t-2}$ and one to three lags of the remaining

[11] The first stage results are reported in Tables 3.4, 3.5, 3.6, and 3.7, under columns OLS. The detailed results for the second stage regressions are available uppon request to the authors.

[12] We implemented the bootstrap procedure following Davidson and MacKinnon (2004). We used the estimated value of $\overrightarrow{y}_{i,t-1}$ as the explanatory variable in the regression instead of the original $\overrightarrow{y}_{i,t-1}$. We substituded missing values of $\overrightarrow{\hat{y}}_{i,t-1}$ for the observed value. In each step of the bootstrap, we substituted zeros for missings in the residuals.

[13] By including only linear terms on neighbouring education, both contemporaneous and lagged, the coefficient of lagged residual is 0.147, with a standard error of 0.084, and a bootstrap $p - value$ of 0.237. If only the lagged neighbouring education is included, this values are 0.145, 0.083, and 0.244, respectively.

Table 3.4: Testing for the presence of country effects, specification I

	OLS	AB91	BB98	SysLX	SysNoCE	LSDVC
	(1)	(2)	(3)	(4)	(5)	(6)
$\overrightarrow{y}_{t-1}$	.924***	.676***	.95***	.92***	.913***	.96***
	(.02)	(.147)	(.103)	(.096)	(.045)	(.036)
$\overrightarrow{S}_t$	.187***	-.049	.24***	.27***	.275***	.091
	(.05)	(.085)	(.093)	(.09)	(.072)	(.058)
$\overrightarrow{S^2}_t$	-.008***	-.0003	-.013**	-.015***	-.015***	-.004
	(.003)	(.005)	(.005)	(.005)	(.004)	(.004)
$\overrightarrow{S}_{t-1}$	-.097**	-.062	-.153**	-.167***	-.175***	-.113*
	(.049)	(.062)	(.066)	(.062)	(.062)	(.06)
$\overrightarrow{S^2}_{t-1}$	.006	.004	.011**	.012***	.013***	.007
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
Obs	300	225	300	300	300	300
Sargan		21.692	44.351	45.434	47.111	
Sargan-*df*		14	25	27	29	
Sargan-*pv*		.085	.01	.015	.018	
DiffS(AB)			22.659	23.742	25.419	
DiffS(AB)-*pv*			.02	.034	.045	
DiffS(P)				1.084	1.676	
DiffS(P)-*pv*				.582	.433	
AR(1)		-2.787	-3.265	-3.242	-3.365	
AR(1)-*pv*		.005	.001	.001	.0008	
AR(2)		.82	1.395	1.391	1.423	
AR(2)-*pv*		.412	.163	.164	.155	
Hausman						19.2
Hausman-*pv*						.014

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%. Standard errors in parentheses. The dependent variable is real Log GDP per worker. DiffS(AB) is the difference Sargan test comparing the model with column (2); DiffS(P) is the difference in the Sargan statistic to the previous column; *df* stands for degrees of freedom; and *pv* stands for $p - value$. AR(1) and AR(2) are tests for first- and second-order serial correlation in first-differenced residuals. All regressions include time dummies, which are always jointly significant. The number of countries for all regressions is 75. Details about the estimations are provided in Section 3.4.2.

Table 3.5: Testing for the presence of country effects, specification II

	OLS	AB91	BB98	SysLX	SysNoCE	LSDVC
	(1)	(2)	(3)	(4)	(5)	(6)
$\overrightarrow{y}_{t-1}$	.902***	.680***	.820***	.850***	.864***	.905***
	(.021)	(.123)	(.070)	(.055)	(.035)	(.047)
$\overrightarrow{S}_t$	.208***	-.046	.351***	.322***	.313***	.094
	(.053)	(.079)	(.093)	(.071)	(.067)	(.062)
$\overrightarrow{S^2}_t$	-.011***	.003	-.020***	-.019***	-.018***	-.005
	(.004)	(.006)	(.006)	(.004)	(.004)	(.004)
$\overrightarrow{S}_{t-1}$	-.116**	-.096	-.217***	-.210***	-.206***	-.163***
	(.051)	(.060)	(.061)	(.057)	(.054)	(.059)
$\overrightarrow{S^2}_{t-1}$	.007*	.006	.015***	.014***	.014***	.010**
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
$G(1)\overrightarrow{S}_t$	-.074	-.476**	-.363	-.120	-.122	-.079
	(.232)	(.231)	(.300)	(.260)	(.267)	(.243)
$G(1)\overrightarrow{S^2}_t$	.008	.014	.032	.007	.008	.005
	(.022)	(.016)	(.026)	(.022)	(.024)	(.022)
$G(1)\overrightarrow{S}_{t-1}$	.086	.375**	.346	.180	.172	.315
	(.208)	(.174)	(.234)	(.234)	(.229)	(.215)
$G(1)\overrightarrow{S^2}_{t-1}$	-.004	-.027*	-.028	-.006	-.006	-.016
	(.022)	(.016)	(.024)	(.022)	(.023)	(.023)
Obs	300	225	300	300	300	300
Sargan		38.976	59.176	61.150	62.590	
Sargan-*df*		26	45	49	51	
Sargan-*pv*		.049	.076	.114	.128	
DiffS(AB)			20.200	22.174	23.614	
DiffS(AB)-*pv*			.383	.510	.542	
DiffS(P)				1.975	1.440	
DiffS(P)-*pv*				.740	.487	
AR(1)		-3.046	-3.160	-3.222	-3.265	
AR(1)-*pv*		.002	.002	.001	.001	
AR(2)		1.030	1.546	1.410	1.433	
AR(2)-*pv*		.303	.122	.159	.152	
Hausman						19.461
Hausman-*pv*						.078
JointWEdu	2.781	3.412	1.239	4.148	4.175	7.391
JointWEdu-*pv*	.027	.013	.302	.004	.004	.117

Notes: See note to Table 3.4. JointWEdu and JointWEdu-*pv* are the test statistic and its $p-value$ for a Wald joint test on the significance of the coefficients of the neighbouring variables. For all regressions we restricted λ to 1.

Table 3.6: Testing for the presence of country effects, specification III

	OLS	AB91	BB98	SysLX	SysNoCE	LSDVC
	(1)	(2)	(3)	(4)	(5)	(6)
$\overrightarrow{y}_{t-1}$	.903***	.719***	.872***	.886***	.879***	.921***
	(.021)	(.135)	(.093)	(.089)	(.045)	(.036)
$\overrightarrow{S}_t$	.202***	-.001	.277***	.286***	.281***	.103*
	(.05)	(.076)	(.075)	(.07)	(.065)	(.059)
$\overrightarrow{S^2}_t$	-.01***	-.002	-.015***	-.016***	-.016***	-.005
	(.003)	(.005)	(.005)	(.004)	(.004)	(.004)
$\overrightarrow{S}_{t-1}$	-.116**	-.09	-.167***	-.186***	-.182***	-.134**
	(.049)	(.057)	(.05)	(.05)	(.051)	(.054)
$\overrightarrow{S^2}_{t-1}$	.007*	.005	.012***	.013***	.012***	.007*
	(.004)	(.003)	(.003)	(.003)	(.003)	(.004)
$G(1)\overrightarrow{S}_t$	.0004	-.154	-.088	-.092	-.068	-.019
	(.062)	(.192)	(.088)	(.08)	(.075)	(.077)
$G(1)\overrightarrow{S}_{t-1}$	.057	.175**	.122**	.163**	.142*	.17**
	(.062)	(.078)	(.058)	(.076)	(.073)	(.076)
Obs	300	225	300	300	300	300
Sargan		26.045	48.752	50.92	52.023	
Sargan-*df*		20	35	38	40	
Sargan-*pv*		.164	.061	.078	.096	
DiffS(AB)			22.708	24.876	25.979	
DiffS(AB)-*pv*			.091	.128	.167	
DiffS(P)				2.168	1.103	
DiffS(P)-*pv*				.538	.576	
AR(1)		-3.087	-3.117	-3.14	-3.242	
AR(1)-*pv*		.002	.002	.002	.001	
AR(2)		.778	1.329	1.277	1.301	
AR(2)-*pv*		.436	.184	.202	.193	
Hausman						18.379
Hausman-*pv*						.049
JointWEdu	5.385	6.243	2.515	4.132	5.049	5.976
JointWEdu-*pv*	.005	.003	.088	.02	.009	.05

Notes: See note to Table 3.4. JointWEdu and JointWEdu-*pv* are the statistic and the respective $p-value$ for a Wald joint test on the significance of the coefficients of the neighbouring variables. For all regressions we restricted λ to 1.

Table 3.7: Testing for the presence of country effects, specification IV

	OLS	AB91	BB98	SysLX	SysNoCE	LSDVC
	(1)	(2)	(3)	(4)	(5)	(6)
$\overrightarrow{y}_{t-1}$	.903***	.677***	.868***	.873***	.875***	.919***
	(.021)	(.129)	(.078)	(.082)	(.042)	(.035)
$\overrightarrow{S}_t$	.202***	-.001	.245***	.286***	.281***	.096
	(.049)	(.073)	(.082)	(.073)	(.067)	(.06)
$\overrightarrow{S^2}_t$	-.01***	-.001	-.013***	-.016***	-.016***	-.004
	(.003)	(.005)	(.005)	(.004)	(.004)	(.004)
$\overrightarrow{S}_{t-1}$	-.116**	-.092*	-.159***	-.182***	-.181***	-.13**
	(.048)	(.055)	(.053)	(.051)	(.051)	(.057)
$\overrightarrow{S^2}_{t-1}$	.007*	.005	.011***	.013***	.013***	.007*
	(.004)	(.003)	(.004)	(.003)	(.003)	(.004)
$G(1)\overrightarrow{S}_{t-1}$	.057***	.193***	.094**	.081***	.079***	.157**
	(.017)	(.063)	(.043)	(.03)	(.025)	(.061)
Obs	300	225	300	300	300	300
Sargan		27.882	50.383	51.848	52.073	
Sargan-*df*		21	36	39	41	
Sargan-*pv*		.144	.056	.082	.115	
DiffS(AB)			22.501	23.966	24.191	
DiffS(AB)-*pv*			.095	.156	.234	
DiffS(P)				1.465	.225	
DiffS(P)-*pv*				.69	.893	
AR(1)		-3.055	-3.185	-3.189	-3.281	
AR(1)-*pv*		.002	.001	.001	.001	
AR(2)		.667	1.291	1.325	1.341	
AR(2)-*pv*		.505	.197	.185	.18	
Hausman						16.089
Hausman-*pv*						.065

Notes: See note to Table 3.4. For all regressions we restricted λ to 1.

variables as instruments in first-difference equations. In columns (3) we present the results for the system estimation. In addition to the previous instruments, we use $\Delta \overrightarrow{y}_{t-1}$ and contemporaneous first-differences of the other regressors as instruments. In column (4), SysLX, we test for the absence of correlation between the education variables and the composite error term in equation (3.9). In this case we assume that these variables are not correlated with the country effect, which would allow us to use its levels as instruments in the level equations of the system estimation. Because lag income, $\overrightarrow{y}_{t-1}$, is still treated as correlated with the country effect, $\overrightarrow{\delta}$, we use $\Delta \overrightarrow{y}_{t-1}$ as instruments in the level equations. In column (5) we implement the full test for the absence of country effects. We reproduce the instruments used in column (4), and $\overrightarrow{y}_{t-3}$ is used as an instrument for income in the level equations. The extra instruments are $\overrightarrow{y}_2$ for period 5, and $\overrightarrow{y}_1$ for period 4. For period 5 we already have $\overrightarrow{y}_4 - \overrightarrow{y}_3$ as an instrument for the level equation. In order to avoid colinearity among the instruments, we used $\overrightarrow{y}_2$ as the extra instrument. Similar reasoning applies to period 4.[14]

The final column in Tables 3.4, 3.5, 3.6, and 3.7 shows the estimation results for equation (3.9) using the LSDVC as presented in Bruno (2005a). We implement a Hausman test that compares the estimates in this model with the ones obtained by OLS. Estimates in column (6) are consistent both under the null and under the alternative, while OLS estimates are only consistent under the null that there are no country effects. It is important to note that, according to Bruno (2005b), the system GMM estimation tends to perform better than LSDVC in the presence of highly persistent series.

Results for specification I, Table 3.4, indicate that while we do not reject the validity of the instruments used for the first-differenced regression, column (2), we do reject their validity for the following 3 estimation alternatives.[15] As such, we are not able to test for the presence of country effects using the GMM procedures applied to this specification. Portela et al. (2004) using a broader set of countries, and imposing a linear relation between education and income, do not reject the absence of country effects. However, under the current specification, using LSDVC and the Hausman test leads to the rejection of the absence of country effects, which contradicts the result obtained from OLS residuals. The mixed results do not allow us to conclude decisively on the absence of country effects in this specification.

Once we include neighbouring education, the different results indicate that we cannot reject the absence of country specific effects (Tables 3.5, 3.6, and 3.7). A possible

[14] The procedure used to test for country specific effects is based on the discussion in Arellano (2003, p.124).

[15] For column (2) the Sargan statistic is 21.7, with a $p-value$ of 0.085; i.e, we do not reject the validity of the instruments at the 1% and 5% significance level. In columns (3) to (5) the $p-value$ associated with the Sargan test is always below 0.018; i.e, the Sargan test rejects the validity of the instruments used in each specific regression. As stated above, we report the corrected standard errors based on the 2–step GMM estimation procedure in columns (3), (4), and (5) in Tables 3.4, 3.5, 3.6, and 3.7.

explanation could be that neighbouring education is, somehow, capturing countries' fixed effect. When we include all four terms for neighbouring education (Table 3.5), both the GMM and the LSDVC estimations corroborate the previous result based on OLS residuals. That is, we cannot reject the null hypothesis that there are no country effects in this specification. Sargan tests do not reject the validity of the instruments used in columns (2) to (5) (in column(2) the test would marginally reject the instruments at the 5% level). By looking to column (5), where we perform the full test for the absence of country effects, we conclude: (i) the instruments are not globally rejected, (ii) the extra-instruments introduced as a departure from column (2), the first-differenced GMM estimation, are also not rejected (the Sargan statistic is 23.6 , with a $p - value$ of 0.542), and (iii) the extra-instruments introduced both in columns (4) and (5) are also not rejected (see DiffS(P) statistics and $p - values$). Finally, the Hausman test performed in column (6) does not reject the validity of OLS estimates in column (1). Although individually the neighbouring coefficients are generically statistically not significant, they tend to be jointly significant, as it is shown by the Wald joint test. The system GMM results shown in column (3) are the exception.

In Table 3.6 we eliminate the nonlinear effects of neighbouring education. For the system estimation, BB98, the Sargan test indicates that we do not reject the instruments used at the 1% and 5% level. However, we would reject them at the 10% level of significance (column 3). Using the difference Sargan test, DiffS(AB), we do not reject the extra instruments used in the system estimation, when compared to the estimation in first-differences. From column (4) we conclude that both the Sargan test and the difference Sargan test do not reject the validity of the instruments used. When we compare with the system estimation (see the value for DiffS(P)) we still do not reject the validity of the new instruments. Results in column (4) have three more instruments compared to those in column (3). This occurs because we can use one more period for the instruments in levels and for the education variables, $\overrightarrow{S}_t$, $\overrightarrow{S^2}_t$, and $G(\lambda)\overrightarrow{S}_t$, respectively. In column (5) the complete set of instruments would be marginally rejected at the 10% level of significance, and the extra instruments are not rejected against both AB91 estimation, and SysLX. We interpret this result as evidence in favour of absence of country effects once we control for $\overrightarrow{y}_{t-1}$. The Hausman test comparing LSDVC and OLS marginally rejects the absence of country effects with a $p - value$ of 0.049. The neighbouring variables are jointly significant across the different estimation alternatives, and the coefficient on lag neighbouring education is statistically significant as well. We then conclude that there is evidence that the short-run effect of neighbours is not significant.

Table 3.7 shows the results when we include only the lagged neighbouring education, in

a linear form. The conclusion on country specific effects is stable, and lagged neighbours' education is statistically significant in all estimations. Based on the evidence provided in this section we conclude that $\overrightarrow{\delta}$ is absent in our empirical model, and implement the analysis in Section 3.4.3 of education spillovers across countries using least squares procedures.

3.4.3 Education spillovers

When looking at Tables 3.5 and 3.6, we observe a consistent result that points to the lack of significancy of contemporaneous neighbouring education. Taking this into account, and the fact that it takes time for spillovers of neighbouring education to occur, we have decided not to use it as a regressor in the following estimations. We restrict $\beta_0 = \beta_{02} = \beta_{12} = 0$ in equation (3.9), and spillovers are identified through the coefficient of $G(\lambda)\overrightarrow{S}_{t-1}$, β_1.

If there are no country effects, endogeneity could still come from correlation of the explanatory variables with $\overrightarrow{\varepsilon}_t$. We first argue that current unobservables that determine countries' income do not interfere with contemporaneous average education in the neighbourhood. The reasoning is that it would take a long time for current perturbations in own income to have an impact on other countries' decisions on education. Within the time frame of our analysis it seems plausible to assume that neighbouring schooling is exogenous. Particularly, we assume that there is no reverse causality from $\overrightarrow{\varepsilon}_t$ to lag neighbouring education, $G(\lambda)\overrightarrow{S}_{t-1}$. The result we achieve in column (4) of Table 3.7 gives strength to an exogeneity argument by not rejecting the exogeneity of the education variables.

Table 3.8 reports the main results for our analysis of education spillovers. In the first three columns neighbouring education is based on the standardized weights as defined by equation (3.6), while in the last three columns we use the non-standardized weights defined by equation (3.7). In columns (1) and (4) we restrict to 1 the decay parameter, λ, while in the remaining columns we estimate its value. This parameter measures how fast distant neighbours become irrelevant to the definition of neighbourhood. The different estimations reported in this table are implemented by nonlinear least squares.

A general conclusion is that, apart from the estimates of λ and β_1, results for the remaining coefficients are relatively stable. By imposing $\lambda = 1$, and using within country standardized weights, column (1), the effect of neighbouring education is positive and significant. When we jointly estimate the decay parameter (see column (2)), we get a smaller coefficient estimate of $G(\lambda)\overrightarrow{S}_{t-1}$, but still statistically significant at the 1% level. A one year increase in neighbouring education has a short-run return of 2.9%.

Table 3.8: Income regressions

	g_{ij}			g^*_{ij}		
	(1)	(2)	(3)	(4)	(5)	(6)
$\overrightarrow{y}_{t-1}$	.903***	.909***	.904***	.902***	.894***	.894***
	(.023)	(.023)	(.024)	(.023)	(.023)	(.023)
$\overrightarrow{S}_t$	.202***	.214***	.212***	.188***	.178***	.176***
	(.053)	(.054)	(.052)	(.051)	(.050)	(.051)
$\overrightarrow{S^2}_t$	-.01***	-.011***	-.01***	-.008***	-.007**	-.007**
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
$\overrightarrow{S}_{t-1}$	-.116**	-.117**	-.123**	-.099**	-.081	-.078
	(.051)	(.052)	(.051)	(.05)	(.05)	(.05)
$\overrightarrow{S^2}_{t-1}$	.007*	.007**	.007**	.006*	.005	.005
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
$G(\lambda)\overrightarrow{S}_{t-1}$	.057***	.029***		.00005***	.0005**	
	(.016)	(.008)		(.00001)	(.0003)	
$\overline{G(\lambda)\overrightarrow{S}_{t-1}}$			.053**			.0006**
			(.02)			(.0003)
$G(\lambda)\overrightarrow{S}'_{t-1}$			.166**			.00005
			(.078)			(.0005)
λ	1	6.051	1***	1	.34**	.319**
		(6.349)	(.191)		(.172)	(.162)
ISR	.064	.062	.068	.074	.069	.068
ILR	.342	.367	.373	.495	.481	.475
ELR	.576	.319	.515	.016	.181	.208
TLR	.918	.686	.888	.512	.661	.683
Half-Life	71.255	76.151	71.921	70.526	65.275	65.458
ImRet	.054	.051	.058	.056	.049	.049
IRet40	.11	.112	.119	.161	.166	.164
TRet40	.296	.209	.284	.166	.229	.236
Obs	300	300	300	300	300	300
Adj-R^2	.962	.963	.962	.964	.964	.964
RMSE	.205	.204	.205	.202	.2	.2
LL	54.486	56.618	55.76	59.137	63.44	63.826

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%. Robust standard errors in parentheses. The dependent variable is real Log GDP per worker. All regressions are estimated by non-linear least squares, and include time dummies, which are always jointly significant. The number of countries for all regressions is 75. RMSE is root mean squared error, and LL stands for log likelihood assuming iid normal errors. ISR, ILR, ELR, and TLR are the internal short-run, internal long-run, external long-run, and total long-run returns to education, respectively. IRet40 and TRet40 are the internal and total returns to education over a 40-year period. The standardized weights g_{ij} are specified as in equation (3.6), while the non-standardized weight g^*_{ij} are formulated as in equation (3.7). $G(\lambda)\overrightarrow{S}'_{t-1} = G(\lambda)\overrightarrow{S}_{t-1} - \overline{G(\lambda)\overrightarrow{S}_{t-1}}$, and $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$ is the within country average of lag neighbouring education. In columns (1) and (4) we restricted λ to 1.

The decay parameter is not statistically different from 0.[16] The value of λ only stretches or compresses, the distribution of weights, not implying the absence of external effects. Column (2) reports a higher log likelihood value, compared to column (1), indicating a better quality fit. The internal short-run return (ISR) to education is 6.2%, whereas the internal long-run (ILR) returns are, approximately, 37%, and total returns to education (TLR) are about 69%.[17] However, the half-life for the long-run returns is 76 years. External returns (ELR) are 32%, indicating that spillovers of education are significant. If we take into account the 40 years span of our data, and compute returns over this period, we conclude that the internal return ($IRet40$) is 11.2%, and the total return ($TRet40$) is 21%. The external return for the 40-year period is 9.8%. By accounting for the 10 year length of each observation period, and reflecting the idea that part of the return to education benefits countries immediately, we would conclude that the immediate return to education is just above 5%.

In column (3) we decomposed the neighbouring education variable into $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$, its mean within countries, and $G(\lambda)\overrightarrow{S}'_{t-1} = G(\lambda)\overrightarrow{S}_{t-1} - \overline{G(\lambda)\overrightarrow{S}_{t-1}}$, the difference to the country mean. The idea is to check wether once we control for a fixed neighbourhood component we still observe a positive effect of neighbouring education. Both variables show a positive and significant coefficient. Having better educated neighbours is good for GDP, and improvements in neighbours' education also have a positive impact on income. The estimated λ is 1, and is statistically different from zero.

In the following three sets of regressions, columns (4), (5) and (6), we define the weights as $g^*_{ij} = \exp(\lambda w_{ij})$. In this case we assume that both country and neighbours specificities in bilateral trade are relevant for the computation of externalities, since the size of the neighbourhood, $\Sigma_j g^*_{ij}$, is not filtered out in the weighting procedure. As above, we conclude that neighbours matter for countries' income. Focusing our attention in column (5), we conclude that the short-run return to education is close to 7%. The decay parameter, λ, is 0.34 and statistically different from 0; i.e., we conclude that weights are more compressed than the original decay value of 1. The internal long-run return to education is approximately 48%, and the estimated total long-run return to education is about 66%. This implies external returns to education of 18% in the long run. It takes 65 years to achieve half value of long-run returns. Within the time frame of our data the total return is approximately 23%, with an internal return of almost 17% and external returns of more than 6%. The fit of this model is the best when compared with the results shown in columns (1) to (4), with a log likelihood of 63.44.

[16] It is also not statistically different from other positive values; its standard error is particularly high.

[17] See Appendix 3.A for an explanation on the computation of the returns to education. When it applies, the computations are based on the median of the variables.

Finally, we replicate in column (6) the decomposition of neighbouring education made in column (3). We observe that average lag neighbouring education, $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$, plays a role in explaining economic performance, while differences to its mean, $G(\lambda)\overrightarrow{S}'_{t-1}$, are not significantly associated with income. At first sight this result reduces the strength of our conclusion that neighbouring education is relevant in explaining economic growth. Indeed, after controlling for a neighbourhood specific component, $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$, differences to the mean in neighbouring education captured by $G(\lambda)\overrightarrow{S}'_{t-1}$ are not relevant for countries' income. It seems that a good neighbourhood is what matters. However, some caution is needed in interpreting this result. Because we are not standardizing the weights, the variable $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$ imbeds both neighbourhood specificities, namely its size, which is kept constant over time, and neighbours average education.[18]

Table 3.9: Specific returns to education

	The Netherlands		The United Kingdom		The United States	
	(2)	(5)	(2)	(5)	(2)	(5)
ILR	0.072	0.303	0.350	0.306	0.093	0.137
ELR	0.309	0.346	0.024	0.409	0.000	0.346
TLR	0.381	0.649	0.374	0.715	0.093	0.483
IRet40	0.022	0.105	0.107	0.106	0.028	0.047
TRet40	0.116	0.225	0.114	0.247	0.028	0.167

Notes: Column numbers correspond to the numbers in Table 3.8. ILR, ELR, and TLR are the internal long-run, external long-run, and total long-run returns to education, respectively. IRet40 and TRet40 are the internal and total returns to education over a 40-year period. The computations are based on countries' specific education in 2000, their weights and the coefficients reported in Table 3.8. See Appendix 3.A for an explanation on the computation of the returns to education.

In order to better understand these results, we now look at the example of three countries, namely The Netherlands, The United Kingdom and The United States. By using standardized weights, and accounting for diminishing returns and neighbouring effects, columns (2), we conclude that the total long-run return to education for the US is just above 9%, which equals the ILR. This result is within the boundaries defined for private returns. It is also in accordance with the finding of Acemoglu and Angrist (2000)

[18] Since $G(\lambda)$ is time constant, $\overline{G(\lambda)\overrightarrow{S}_{t-1}}$ can be written as $G(\lambda)\overline{\overrightarrow{S}_{t-1}}$, where the elements of $\overline{\overrightarrow{S}_{t-1}}$ are averages of education within countries and over the observation period.

of not significant externalities to education. However, within this specific formulation of the weights, our findings reveal that Acemoglu and Angrist's results cannot be extended to all countries. Even though external returns within the US might be negligible, and the US themselves benefit little from neighbours' education, we conclude that on average the rest of the world does benefit from neighbours' education.

The US is located at the economic and technological frontier, and on the downside of the returns to education. According to our computations based on column (2), Table 3.8, decreasing short-run returns can occur for countries with an average education above 9.7 years.[19] In 2000, the US has an average schooling of 12 years, and shows virtually no long-run external returns to education. The UK has, for the same year, an average education of 9.4 years, an internal long-run return to education of 35%, and an external long-run return of 2.4%. The Netherlands, with 9.4 average years of schooling shows, in the long run, an internal return of 7.2%, and an external return of 31%.

When we do not standardize the weights we observe a substantial increase in returns to $G(\lambda)\overrightarrow{S}_{t-1}$ for this set of countries. This result can be explained by the role each country plays in international trade. In columns (5) we use absolute weights to compute the returns, and this way we are accounting for the specific role of countries in international trade, τ_i and τ_j, *per se* and not only for the relative impact of each country in a specific neighbourhood. In particular, for the computation of the TLR, we add all absolute weights that each country attaches to its neighbours; i.e., we compute the absolute importance of the country-specific neigbourhood. The Netherlands, the UK and the US, as key partners in global trade, have a high weight in international trade, and they trade with important partners.

3.5 Concluding remarks

We find diminishing returns to own education, and a positive role for neighbouring education. Using standardized weights, we conclude that both internal and external returns to education exceed the private returns identified in the literature. If we account for the time interval of the data available, our results are once again compatible with both externalities and the private returns observed. In a 40–year interval, the median overall return is below 30%, with internal returns to schooling being bounded between 11% and 17%. Once we estimate the decay parameter, long-run returns to neighbouring education can be at least as high as 18%, with the half-life period above 65 years. These findings are compatible with the result by Moretti (2004a) that external returns are higher for less

[19] This value follows from: $0.214/(2 \times 0.011) = 9.7$.

educated groups.

Our results reconcile the evidence from cross-country analysis with the empirical findings on private returns. For example, Harmon et al. (2003) conclude that for the UK private returns to education are bounded between 7% and 9% when using OLS, and 11% to 15% in case IV is used. Our computations indicate that for the UK internal returns over a 40 year period would be 11%.

The size of the g_{ij} gives an indication of the size of the optimal decision making unit that internalizes most of the externalities of extra investment in human capital. An area of the size of the US captures 49% of the total weight for itself, whereas an area like the Netherlands captures only 11%. This implies that areas of centralized policy determination of the size of the European Union (EU) or the US would internalize only half of the spillovers. It is hard to believe that policy making can be centralized to such an extent.

This chapter offers an interpretation of the correlation between GDP and average years of education of the workforce as observed in the cross country panel data, which makes sense from an economic point of view. There is a clear distinction between private and social returns (or spillovers). Private returns are by definition local, and therefore confined to the own country. Private returns are immediate (they raise worker's direct productivity), and more or less constant over time (they raise worker's productivity for the rest of his life), so that the long and short-run return are equal. These returns square reasonably well with the Mincerian rate of return; otherwise individuals would have invested in more human capital, at least in countries with well established capital markets and government institutions to facilitate investments in human capital. Finally, there are declining marginal returns to education, so that the return is higher in countries with lower average levels of education. Social returns exceed by definition the borderlines of a country, since there is no reason why spillovers would be confined to the own country. By facilitating new invention, knowledge spillovers affect more the growth than the level of productivity. Their effect is stronger in the long run than in the short run. All these features are born by the data, and all together provide a reasonable description of the data generating process.

If one believes this story, one has to come up with an explanation why Acemoglu and Angrist (2000) do not find significant human capital externalities. The main reason is probably in the timing and the place. The instruments that they use are probably well suited to detect changes that are implemented in a short span of time and that are highly localized. They are probably less suitable for changes that take decades to materialize and that affect areas of the size of the US or the EU. Our regressions suggest that it is the case

with human capital spillovers. It takes about seven decades for half the spillovers to be realized, and the effects are spread out over wide regions. For these cases, our regression perform better, though we immediately admit that it takes a long time series to establish the dynamics of a process with such a long half time, a much longer time series than is available at present, given that the education revolution started just some 70 years ago.

Appendices to Chapter 3

3.A Computation of returns to education

Computation of the different returns to education is based on equation (3.9) with the restriction $\beta_0 = \beta_{02} = \beta_{12} = 0$, and is implemented as follows. The internal short-run return to education (ISR) is defined as

$$ISR = \frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} = \alpha_0 + 2\alpha_{02}\overrightarrow{S}_t \tag{3.11}$$

The internal long-run return (ILR) due to a change of education in period t is defined as

$$\begin{aligned} ILR &= \frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} + \frac{\partial \overrightarrow{y}_{t+1}}{\partial \overrightarrow{S}_t} + \frac{\partial \overrightarrow{y}_{t+2}}{\partial \overrightarrow{S}_t} + ... = \\ &= \frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} + \left(\rho\frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} + \alpha_1 + 2\alpha_{12}\overrightarrow{S}_t + \beta_1\varpi_{ii}\right) + \rho\frac{\partial \overrightarrow{y}_{t+1}}{\partial \overrightarrow{S}_t} + ... = \\ &= \frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} + \left(\rho\frac{\partial \overrightarrow{y}_t}{\partial \overrightarrow{S}_t} + \alpha_1 + 2\alpha_{12}\overrightarrow{S}_t + \beta_1\varpi_{ii}\right)\frac{1-\rho^\infty}{1-\rho} = \\ &= \frac{\alpha_0 + \alpha_1 + 2\left(\alpha_{02} + \alpha_{12}\right)\overrightarrow{S}_t + \beta_1\varpi_{ii}}{1-\rho} \end{aligned} \tag{3.12}$$

where $\varpi_{ii} = \frac{\partial G(\lambda)\overrightarrow{S}_t}{\partial \overrightarrow{S}_t}$. Taking the derivative of neighbouring education in order to own education we obtain the weight country i attaches to itself.

The total long-run return (TLR) is given by

$$TLR = \frac{\alpha_0 + \alpha_1 + \beta_1 G(\lambda)\iota' + 2\left(\alpha_{02} + \alpha_{12}\right)\overrightarrow{S}_t}{1-\rho} \tag{3.13}$$

We compute TLR under the assumption that neighbouring education, including own country education, changes by one year. When we use standardized weights, by definition, $G(\lambda)\iota'$ simplifies to a vector of ones. Using non-standardized weights implies that $G(\lambda)\iota'$ contains the sum of absolute weights each country attaches to his neighbours. Comparing equations (3.12) and (3.13), we exclude $\beta_1\varpi_{ii}$ to avoid double counting. Finally, the external long-run return (ELR) is computed as $ELR = TLR - ILR$.

The short-run return can be generically defined as $s = l\left(1 - e^{rt}\right)$, where l is the long-run return, and r is the annual rate of convergence. If we take a first order Taylor expan-

sion of the function $1-e^{rt}$ around $t = 0$, we obtain rt. This function defines the immediate return as a fraction of the long-run return. In our case, $r = (1-\rho)/10$, and the time needed to achieve the immediate return, t^*, is defined as $ISR = ILR\left(1 - e^{-(1-\rho)/10\times t^*}\right)$. By using data for periods of 10 years we measure t^* imperfectly. We assume that innovations in education are distributed uniformly, and correct t^* by subtracting half of the interval, 5 years. The immediate return is then defined as $ImRet = r \times (t^* - 5) \times l$.

For the 40–year interval, the return to education is defined as

$$IRet40 = ILR \times \left(1 - e^{-(1-\rho)/10\times 40}\right) \tag{3.14}$$

and similarly for $TRet40$.

3.B Data description

Table 3.10: List with the 75 countries used in the analysis

Argentina	Ecuador	Italy	New Zealand	Sweden
Australia	El Salvador	Jamaica	Nicaragua	Switzerland
Austria	Finland	Japan	Niger	Syria
Bangladesh	France	Jordan	Norway	Tanzania
Barbados	Ghana	Kenya	Pakistan	Thailand
Belgium	Greece	Korea, South	Panama	Togo
Bolivia	Guatemala	Lesotho	Paraguay	Trinidad and Tobago
Brazil	Honduras	Malawi	Peru	Turkey
Cameroon	Hong Kong	Malaysia	Philippines	Uganda
Canada	Iceland	Mali	Portugal	United Kingdom
Chile	India	Mauritius	Romania	United States of America
Colombia	Indonesia	Mexico	Senegal	Uruguay
Costa Rica	Iran	Mozambique	South Africa	Venezuela
Denmark	Ireland	Nepal	Spain	Zambia
Dominican Republic	Israel	Netherlands	Sri Lanka	Zimbabwe

Table 3.11: Sample of weights for The Netherlands

Country	Partner	$e^{w_{ij}}$	g_{ij}
Netherlands	France	38.447	0.175
Netherlands	Belgium	35.140	0.160
Netherlands	Italy	29.717	0.135
Netherlands	Netherlands	24.243	0.110
Netherlands	United Kingdom	16.369	0.074
Netherlands	Spain	10.219	0.046
Netherlands	Portugal	5.810	0.026
Netherlands	Iran	5.470	0.025
Netherlands	Brazil	5.295	0.024
Netherlands	Turkey	4.546	0.021

Note: We restricted λ to 1.

Table 3.12: Sample of weights for The United States

Country	Partner	$e^{w_{ij}}$	g_{ij}
United States of America	United States of America	126.096	0.490
United States of America	France	13.121	0.051
United States of America	Italy	11.729	0.046
United States of America	Mexico	9.253	0.036
United States of America	Japan	8.853	0.034
United States of America	Brazil	8.849	0.034
United States of America	Indonesia	6.389	0.025
United States of America	India	5.993	0.023
United States of America	United Kingdom	5.769	0.022
United States of America	Canada	5.397	0.021

Note: We restricted λ to 1.

Chapter 4

Gender segregation and the wage gap in Portugal: an analysis at the establishment level

4.1 Introduction

The composition of the labour force differs widely across employers. Two main lines of reasoning have been followed to explain that pattern: taste-based or quality-sorting recruitment. In the first case, preferences by employers (or co-workers or customers) will lead an employer into recruiting particular types of workers, but not others. Becker (1971) has set the stage for this analysis, under the heading discrimination in the labour market. The other line of reasoning distinguishes workers by their *quality* or productivity, to stress sorting effects, according to which similar workers will be matched together in firms, if their skills are complements in the production process. A good version of this type of model is presented in Kremer (1993) and Kremer and Maskin (1996). Both theories predict that workers with different attributes will be segregated into different workplaces.

Employment segregation will be a source of wage differentials between groups of workers to the extent that different firms pay different wages. The two theories mentioned diverge, however, on the implications of segregation for wage setting. Nevertheless, gender segregation along occupation or industry lines has been subject to wider scrutiny than gender segregation among establishments. Studies evaluating the impact of the degree of femaleness of the establishment on wages have in general found that inter-establishment gender segregation accounts for a substantial share of the wage gap (see Carrington and Troske, 1995, 1998; Yoon et al., 2003; Reilly and Wirjanto, 1999; Groshen, 1991; Pfeffer and Davis-Blake, 1987; and, for earlier awareness on this pattern, McNulty, 1967, and Buckley,1971).

This chapter aims at quantifying the trend in worker segregation across establishments and its impact on wages in Portugal over a fifteen year period. We concentrate on the gender dimension, to answer the questions: What is the level of gender segregation across establishments in the Portuguese labour market and how has it evolved over time? What is the impact of segregation on wages? Is that impact different for men and women? The aim is also to contribute to a better understanding of the Portuguese gender wage gap, which revealed a hump-shaped pattern from 1985 to 1999, reaching a peak in 1991.

The study relies on a large linked employer-employee data set gathered by the Ministry of Employment, based on an inquiry that every firm with wage-earners is legally obliged to fill in. Each year, an average of two million workers, 200 thousand establishments, and 150 thousand firms are covered.

We evaluate worker segregation across establishments as departures from the segregation that would prevail if workers were randomly assigned to establishments, instead of departures from perfect integration, if groups were proportionately represented in each establishment. In fact, Carrington and Troske (1997) have proven that, in particular in the presence of small units, different groups of workers will never be evenly distributed across establishments, even if the allocation is determined on a random basis. This result implies the computation of systematic segregation, as opposed to simple segregation, being an important improvement to the literature. We therefore compute random and systematic segregation, using both the Gini and the dissimilarity indices. The impact of the degree of femaleness of the establishment on wages will be quantified using the Oaxaca and Cotton-Neumark procedures for wage decomposition.

The chapter is organized as follows. Section 4.2 describes the data set. Section 4.3 provides information on the Portuguese labour market and describes trends in female employment. Section 4.4 analyses systematic gender segregation across establishments, and Section 4.5 discusses the impact of gender segregation on wages. Concluding comments are presented in Section 4.6.

4.2 Data set

The data used, *Quadros de Pessoal*, are gathered annually by the Ministry of Employment, based on an inquiry to every firm with wage-earners, which reports information on the firm, the establishment and all of its workforce. Given the legally binding nature of the inquiry, the response rate is extremely high. The fact that the information (namely on wages) is provided by the employer and the legal request for the data to be permanently displayed in a public space in the establishment, contribute to their reliability, reducing

measurement errors. Reported data include the worker's gender, age, skill, occupation, schooling, tenure, earnings and duration of work, and the establishment's and the firm's location, industry, and size.

The full coverage of the workforce within establishments is a clear advantage of this data set for the study of worker segregation across establishments. Also, the data are very representative, being in fact a census of the establishments employing paid labour. A wide set of variables is reported for each worker, but nevertheless less rich information is provided on establishments.

Establishments in manufacturing and the services have been kept for analysis.[1] Only wage-earners aged 16 to 65 were considered. Note that a minimum establishment size requirement had to be imposed for the analysis of the homogeneity of the workforce within firms. Indeed, it would be meaningless to compute segregation for an establishment with one worker or a similarly tiny dimension. During the period under analysis, the legislation in the country defined a *micro-firm* as one employing less than 5 workers. We have considered that benchmark, and throughout the chapter the analysis is restricted to establishments with at least 5 workers. Therefore, large and small firms —which may be different in terms of work organization and labour flexibility, for example —are included in the analysis, and just tiny ones have been dropped.

These criteria led to a data set of 1.4 million workers and 62 thousand establishments on average each year. A high proportion of the establishments in the Portuguese economy has less than 5 wage-earners, but nevertheless they employ a very small proportion of the workforce. Indeed, 90 percent of the wage-earners in the selected industries and age bracket is kept for analysis, even though just 40 percent of the establishments employing them fulfill the size requirement.[2]

4.3 The Portuguese labour market and trends in female employment

Interest in the Portuguese labour market has widened over the last two decades, mainly driven by its good performance after mid-1980s, when compared to other Western econo-

[1] Given the low share of wage-earners in agriculture, its coverage in the data set is low. Because of that we have excluded it from the analysis. By using manufacturing and services in our analysis we follow the practice in the literature, and report comparable results.

[2] We have also considered establishments employing at least 3 workers, and performed the computation of the overall segregation index. The level of total segregation is slightly higher once those establishments are included in the analysis, whereas its trend is remarkably similar to the one that will be reported in Section 4.4.

mies. For example, the unemployment rate declined from 9 percent in 1985 to 4 percent in 2000. The economy has been under a process of modernization and restructuring, mainly after joining the European Union in 1986.

As a consequence of this process of change, demand for skilled workers increased and overall wage inequality widened, as wages at the top of the distribution grew faster than those at the lower end (Cardoso, 1998). This has been associated with rising returns to education and job requirements, and is consistent with the skill-biased technological change hypothesis (see Hartog et al., 2001). In addition, wage differentials associated namely with firm size and industry affiliation are substantial when compared with other European countries (see for instance Hartog et al., 2000). In particular, the size of inter-industry wage dispersion is high, comparable to countries normally rated as having a decentralized wage setting system, such as the United States (US) or Canada.

These large wage differences for apparently equally-skilled workers indicate flexibility to exploit industry or firm and establishment specific conditions, which may be related to particular circumstances regarding industrial relations. Indeed, high wage flexibility has been pointed out as a particular feature of this market (OECD, 1994a), and studies at the micro level have shown that firms have considerable degree of freedom when manipulating wages, despite widespread collective bargaining (Cardoso and Portugal, 2005).

In addition, it is well-known that wages are in Portugal low compared to other Western economies. These lower wages reflect lower productivity of labour, which itself may indicate reduced levels of physical and human capital (see Branco and Mello, 1992).

Female participation in the labour market is high when compared to other Southern European countries such as Spain, Italy and Greece, and above the European Union average. Female employment has been steadily increasing in the Portuguese economy. Whereas it accounted for 32 percent of total employment in manufacturing and the services in 1985, by 1999 it had risen to 43 percent. The composition of female employment underwent changes as well. The share of employed women holding a University diploma increased during that period from 3 percent to 9 percent, while the share holding a High-School diploma increased from 11 percent to 19 percent. Changes in the composition of male employment have been slower, as the share of employed males holding a University diploma increased from 4 percent to 7 percent, and the share holding a High-School diploma increased from 11 percent to 16 percent. These values illustrate clearly the low level of educational attainment of the working population in Portugal.

Raw data points to a certain degree of gender segregation at the establishment level (see Tables 4.4 and 4.5 in Appendix 4.A). While in the sample of females the average share of women in the establishment is 56 percent to 65 percent, in the sample of males

the values range from 20 percent to 25 percent —females tend to have predominantly female co-workers, and males tend to have predominantly male co-workers.

Economic growth and increasing integration of women into the labour market did not lead in Portugal to a systematic decline in the gender pay gap. In fact, the gap measured as the difference between the mean values of log-wages in each group increased from 1985 to 1991, declining afterwards. Furthermore, empirical evidence has shown that even after controlling for several worker and employer attributes, the Portuguese wage gap is high and persistent (Kiker and Santos, 1991; Vieira, 1999).

4.4 Gender segregation at the establishment level: systematic and random components

To evaluate total segregation in the labour force, the Gini and the Duncan dissimilarity indices, respectively G and D, have been used:

$$D = \sum_{i=1}^{T} \frac{1}{2} |w_i - m_i| \tag{4.1}$$

where w_i and m_i are the establishment i's share of female and male employees in the sample, respectively, and T is the number of establishments in the sample, and:

$$G = 1 - \sum_{i=1}^{T} w_i \left(m_i + 2 \sum_{j=i+1}^{T} m_j \right) \tag{4.2}$$

with the calculation being performed in the data sorted by w_i/m_i. Both indices are bounded between 0 and 1, with 0 corresponding to maximum evenness (perfect integration), and 1 to maximum unevenness.

In intuitive terms, the value of the Duncan index indicates the share of men (or women) that would have to move to eliminate inter-firm segregation (see Carrington and Troske, 1995, p.517). Thinking in terms of segregation curves, "the Gini index is equal to two times the area between the diagonal line and the segregation curve, while the dissimilarity index is equal to the maximum vertical distance between the diagonal and the segregation curve" (Hutchens, 2001, p.17). For a more extensive discussion of the interpretation of the indices, see Flückiger and Silber (1999, p.53-62).

Hutchens (2001) provides a thorough discussion of the properties of segregation indices. Out of seven desirable properties highlighted, the Gini index fails to meet additive decomposability (i.e., if we partition the population into mutually exclusive groups, the

total Gini index cannot be exactly decomposed into the between-groups plus the within-groups components), whereas the dissimilarity index fails to meet the property relating to the movement of individuals between groups (for example, if women were shifted from a group with lower proportion of women to one with higher proportion of women, the dissimilarity index could nevertheless decrease). Hutchens himself does not attach much relevance to the problem of the Gini index, acknowledging that this is a useful measure. That is particularly so in our case, since we are not interested in knowing the contribution of a subset of establishments to total segregation. By verifying the remaining six properties, the Gini index allows for the ordering, in a credible way, of different distributions in terms of their level of segregation. The problem with the dissimilarity index is potentially more serious. However, that index has been extensively used in the previous literature (see for example Carrington and Troske, 1995, 1997, 1998; or Yoon et al., 2003) and, if we want to have benchmark results to compare with ours, we are bound to use the same type of segregation indices. Or, as Hutchens acknowledges, "the dissimilarity index and the Gini index dominate the empirical literature" (Hutchens, 2001, p.17) and we have therefore chosen to use them.

Segregation will never reach the value 0, in particular if the economy is made up of small units, even if workers are randomly allocated to establishments. The example in Carrington and Troske (1998, p.450-451) helps to clarify this point: in an economy made up of a large number of two-worker establishments, whose labour force is assigned randomly from a population with an equal number of men and women, one would end up with one quarter of the establishments with two men, one quarter with two women, and half with one man and one woman. This would imply a Gini segregation index of 0.75, and a Dissimilarity index of 0.5. A generalization of this result for different proportions of females in the labour force and different classes of establishment size is provided in Carrington and Troske (1997, p.403-404; 1998, p.451), showing that random allocation generates positive segregation as measured by traditional indices, and that reported segregation increases when the sample is made up of smaller establishments, therefore rendering comparisons across samples non-trivial.

One should therefore quantify the degree of systematic segregation existing in the sample evaluated as departures from random segregation (the one that would result from pure chance in the allocation of workers to establishments), instead of departures from absolute evenness. This idea was discussed and applied in Boisso et al. (1994), as well as in Carrington and Troske (1997, 1998).

To compute random segregation, we consider the original number of females and males and the original establishment sizes in the sample. Then, workers are randomly reallocated

to establishments and the segregation indices are computed.[3] After a certain number of replications of this procedure, the average segregation index reached is the random segregation. To obtain the standard errors of the indices (total, random and systematic), we use the bootstrap technique applied to segregation measurement as explained in Boisso et al. (1994) and later also applied by Carrington and Troske (1998). In our computations the bootstrap is based on 100 samples of 10 percent drawn from the original data.[4]

The systematic Gini segregation coefficient is computed as follows (Carrington and Troske, 1997):

$$\hat{G} = \begin{cases} \frac{G-G^*}{1-G^*} \text{ if } G \geq G^* \\ \frac{G-G^*}{G^*} \text{ if } G < G^* \end{cases}, \tag{4.3}$$

where $\hat{G} \in [-1, 1]$. If actual segregation exceeds random segregation ($G > G^*$), then $\hat{G}$ quantifies excess segregation over that dictated by chance, expressed in percentage of the maximum segregation that could occur ($1 - G^*$). When $G < G^*$, we face a situation in which there is excess evenness (Carrington and Troske, 1997, p.406) in the distribution of gender across establishments, that is, not even random allocation would be able to obtain that balance in the distribution of individuals. As this index assesses random deviation, its interpretation is not based on the quota of minorities nor on the size of the units. However, as the size of units increases, the modified segregation index, $\hat{G}$, tends towards the value of the original index, G. The same procedure applies to the dissimilarity index.

Gender segregation across establishments in the Portuguese labour market is high and has been relatively stable between 1985 and 1999 (see Table 4.1). We observe a slight increase in the random segregation, which can be explained by the change in the dimension of establishments[5] and in the female participation in the labour market.

Systematic segregation, when measured by the Gini coefficient, has been stable around 0.67 during this period. The Dissimilarity index reveals as well stability, around the value 0.49. This means that approximately 49 percent of women or men would have to switch employer to come to an equal (random) distribution of gender across establishments. This suggests a high level of segregation when compared to the US manufacturing, since

[3] We use the uniform distribution to generate random numbers that sort workers, before they are matched to the original array of employers (keeping the original number of positions available in each employer). Using a random number generator, we guarantee that the reallocation does not follow a systematic pattern. The procedure used also guarantees that the data set has exactly the original structure (establishment size and gender composition of the workforce).

[4] We have repeated the procedure drawing 200 or 50 samples, and results remained roughly unchanged. We have also checked whether dealing with a sample of the workforce, as most authors are constrained to do, instead of the full population, would influence the results. Also in this case, results change very little. The full set of results is available from the authors upon request.

[5] The average establishment size in the population under study decreased from 28 to 20 workers over the period.

Table 4.1: Gender segregation at the establishment level

	Total Segregation		Random Segregation		Systematic Segregation	
	Dissimilarity	Gini	Dissimilarity	Gini	Dissimilarity	Gini
1985	0.553	0.732	0.121	0.190	0.492	0.670
	(0.016)	(0.016)	(0.005)	(0.007)	(0.017)	(0.019)
1987	0.552	0.737	0.123	0.193	0.489	0.674
	(0.016)	(0.014)	(0.005)	(0.006)	(0.018)	(0.016)
1989	0.556	0.739	0.126	0.197	0.491	0.674
	(0.016)	(0.012)	(0.004)	(0.006)	(0.017)	(0.014)
1991	0.553	0.736	0.129	0.200	0.487	0.670
	(0.014)	(0.011)	(0.004)	(0.005)	(0.015)	(0.014)
1993	0.548	0.733	0.135	0.210	0.478	0.662
	(0.012)	(0.012)	(0.004)	(0.005)	(0.012)	(0.014)
1995	0.559	0.741	0.138	0.214	0.488	0.670
	(0.012)	(0.009)	(0.006)	(0.005)	(0.012)	(0.011)
1997	0.564	0.744	0.141	0.218	0.493	0.672
	(0.009)	(0.010)	(0.004)	(0.005)	(0.010)	(0.013)
1999	0.563	0.742	0.144	0.223	0.489	0.668
	(0.009)	(0.007)	(0.004)	(0.006)	(0.010)	(0.009)

Note: Bootstrap standard errors in parentheses.

Carrington and Troske (1998) have reported values of 0.33 and 0.45, respectively for the dissimilarity and the Gini index. The values for Portugal are, however, remarkably in line with those presented for Korea by Yoon et al. (2003), with an industry coverage similar to ours.

4.5 The impact of gender segregation on wages

To analyse the impact of gender segregation at the establishment level on wages, consider regressions of the type:

$$W_{gi} = \beta_g X_{gi} + \eta_{gi} \tag{4.4}$$

where subscript $g = (m, f)$ indicates the gender, W_{gi} is the log hourly wage of worker i, X_{gi} denotes a set of individual and job related characteristics, which includes the proportion of females in the establishment; β_g denotes the regression coefficients and η_{gi} is a random error term assumed to satisfy the usual properties. Hourly wages were computed as monthly wages divided by the number of hours worked. Tables 4.4 and 4.5 in Appendix 4.A list all the variables included and their descriptive statistics.

From OLS estimation of equations (4.4) it follows that:

$$\bar{W}_m - \bar{W}_f = (\bar{X}_m - \bar{X}_f)\hat{\beta}_m + (\hat{\beta}_m - \hat{\beta}_f)\bar{X}_f \tag{4.5}$$

which is the Oaxaca (1973) male-based decomposition. The first term on the right hand side indicates the portion of the wage gap attributable to differences between sexes in the mean values of productive and job related characteristics (i.e., differences in endowments); the second term represents the portion attributable to differences in prices of those characteristics (often referred to as wage discrimination). The idea of the first term is to value the difference in endowments at the wage rate that would prevail in the economy in the absence of wage discrimination (the non-discriminatory wage structure, following the reasoning by Becker, 1971). Oaxaca suggested using alternatively male or female wages as that reference wage distribution, to define a range within which the values of the components would fall.

Cotton (1988) and Neumark (1988) choose instead the computation of a specific point within that range, by considering the non-discriminatory wage structure ($\hat{\beta}^*$) as the weighted average of the female and male wage structures, with weights equal to their employment shares. The wage decomposition would therefore be defined as follows:

$$\bar{W}_m - \bar{W}_f = (\bar{X}_m - \bar{X}_f)\hat{\beta}^* + (\hat{\beta}_m - \hat{\beta}^*)\bar{X}_m + (\hat{\beta}^* - \hat{\beta}_f)\bar{X}_f \tag{4.6}$$

Differing from Oaxaca's proposal, the last two terms measure the male advantage and the female disadvantage in coefficients (i.e., the extent to which the returns to productive and other characteristics differ from the non-discriminatory returns). These two terms are then used as measures of the extent of labour market discrimination.

It therefore follows that the contribution of the proportion of female workers at the establishment level (P) to the gender gap is given, under the Oaxaca method, by

$$(\bar{P}_m - \bar{P}_f)\hat{\beta}_{mP} + (\hat{\beta}_{mP} - \hat{\beta}_{fP})\bar{P}_f \tag{4.7}$$

and by

$$(\bar{P}_m - \bar{P}_f)\hat{\beta}^* + (\hat{\beta}_{mP} - \hat{\beta}^*)\bar{P}_m + (\hat{\beta}^* - \hat{\beta}_{fP})\bar{P}_f \tag{4.8}$$

under the Cotton-Neumark approach.

4.5.1 Higher concentration of women in the establishment: lower wages for women, but higher for men

The proportion of females in the establishment workforce has a negative impact on females' wages, with the coefficient being statistically different from zero in every year. Conversely, the higher the proportion of females in the establishment, the higher males' wages (except in 1999) (see Tables 4.6 and 4.7 in Appendix 4.A). For example, for males in 1985 an increase of 10 percentage points in the proportion of females in the establishment was associated with an increase of 0.3 percent in the average wage; this coefficient remained relatively stable over the sample period, with a slight decrease after 1995. On the other hand, the negative impact of this variable on female wages increased until early-1990s; by 1993, a 10 percentage point increase in the proportion of females in the establishment was associated with a decline in average female wages of approximately 1 percent. These results contrast to previous available evidence that had revealed that the femaleness of the establishment depressed the wages of both men and women (see Carrington and Troske, 1998; or Reilly and Wirjanto, 1999).

We have checked whether these results could be driven by the aggregate occupational controls used in the regressions. That is not the case, since results are very stable once we include more detailed controls (two- or three-digit classification of occupations). Moreover, the results are also invariant to the exclusion of part-timers from the regression.

Still one other problem might affect the results. Even though the previous literature on this issue has relied on OLS estimation, the endogeneity of the variable share of females might bias the results. We have extensively searched for ways to account for the endogeneity of this variable.[6] Omitted variables is the most likely source of bias in our OLS regressions. We would like to control for firm attributes that may be correlated with the share of females in its establishments, but data limitations force us to include such variables in the error term. Firms may select the share of females they hire based on certain variables we are not controlling for. In particular, we have not controlled for the establishment productivity and it seems reasonable to assume that there are some unobserved productivity differentials captured in the error term of the regression that may be correlated with the share of females included in the equation. For instance, firms with low productivity might tend to employ more female workers because they fit the jobs better,

[6] First of all, we have searched for feasible instruments. The share of female in the occupation or in the region seemed at first sight natural candidates. However, the share of females in the occupation has been used in the literature as a direct determinant of individual wages (see for example Bayard et al., 2003), and the share of females at the regional level can itself be considered a determinant of the wage level of the worker. We concluded that we were unable to find in our data set feasible instruments for the share of females in the establishment.

and females might be less productive than males due for example to job career interruptions. If such a sorting process exists, it would show up in the regression coefficient of the share of females, since productivity is not controlled for.

We believe, however, that this selection issue can be tackled if we include a firm-specific effect in the wage regression. We have thus re-run the wage regressions including a set of firm dummy variables. These terms are bound to capture the heterogeneity across firms in terms of, for example, productivity, product market conditions or average labour quality. Controlling for firm unobservable attributes should, at least partly, account for productivity differences across establishments.

The results for these estimations are reported in Tables 4.8 and 4.9. The results for the male working population using fixed-effects are in line with the ones previously obtained. There is still a positive impact of the share of females in the establishment on male wages. Indeed, from 1985 to 1995 that effect is now much stronger than previously estimated. If the mechanism of sorting based on productivity described above were indeed at work, we would expect these results (since, without any kind of control for the firm productivity, the negative covariance between the share of females and the firm productivity would bias the coefficient on the share of females downwards).[7]

For the female working population, results from 1993 onwards using fixed-effects are consistent with the ones previously obtained using OLS; i.e., a larger share of females in the establishment lowers female wages. That effect is now stronger. However, from 1985 to 1991 we find using fixed effects that a larger share of female had a positive impact also on female wages (as opposed to the previous results using OLS).[8]

The taste-based wage discrimination and the quality-sorting theories reach different predictions regarding wage gaps. According to the sorting theory, the wages of different groups of workers within a firm will be positively correlated (see the matching models in Kremer, 1993; and, for a more general model, Kremer and Maskin, 1996). The wage discrimination theory, on the other hand, allows for the wages of men working with women to be higher than the wages of other men, to compensate them for the 'disutility' of having female co-workers. The evidence that a higher proportion of females in the establishment lowers wages for women but raises wages for men would therefore lend support to wage discrimination type of models. However, comparison of the OLS results with the fixed-effects results highlights the relevance of sorting type of theories for the explanation of

[7] With no control for firm productivity, we would have $E(\hat{b}) = b + c\frac{Cov(P,F)}{var(P)}$, where P stands for the proportion of females, F is the firm productivity, b is the coefficient on the proportion of females and c is the coefficient on productivity, if it were observable and had been included in the regression. Since $c > 0$ and $Cov(P, F) < 0$, the lack of control for productivity biases b downwards.

[8] As expected, the coefficient on the size of the establishment is now considerably lower than before, since most of that effect is absorbed by the firm dummy variable.

the pattern and trend of gender wage setting in Portugal.

Over time, the positive impact of the share of females at the establishment on male wages declined in Portugal. As the proportion of female workers in the economy increased, the compensation that male workers seem to receive for working with females has declined. This result points to a fading out of discrimination mechanisms.

4.5.2 Segregation and the wage gap

We return to the OLS results previously used in the literature (and will comment below on the results of the fixed-effects model). The contribution of the concentration of females at the establishment level to the gender wage differential is quite significant, varying from 10 percent in 1995 to 25 percent in 1989 (see the last column in Table 4.2).

The role of prices has been prominent (see Table 4.2). The Oaxaca methodology using male wages as the benchmark indicates that, concerning the proportion of females at the establishment level, P, the contribution of the endowment component is negative (except in 1999). In fact, given that the share of females has a positive impact on males wages (the reference wage distribution considered) and that women on average work in establishments with a higher proportion of females, the endowment component would raise female wages, reducing the gender wage gap. However, this is offset by the effect of differences in prices (i.e., coefficients) associated with the femaleness of the establishment (precisely because they are positive for men and negative for women, as previously reported). This price component accounts for 15 percent of the observed gap in 1985 and 21 percent in 1999, fluctuating during the period in-between (see Table 4.3).

The decomposition based on the Cotton-Neumark methodology reveals that, for the group of all the variables, differences in endowments, the male advantage and the female disadvantage contribute positively to the observed gender gap, which is in line with the results of Gyimah-Brempong et al. (1992). The contribution of the female disadvantage is larger than the contribution of the male advantage.

With respect to the proportion of females in the establishment, most of the observed gender gap is due to the female disadvantage component, rather than to the male advantage or to differences in endowments, whose contributions to the gap are fairly low. Indeed, female underpayment accounts for 10 percent to 19 percent of the gender pay gap, whereas male overpayment accounts for 2 percent to 4 percent of that gap. This finding is at odds with the results of Rilley and Wirjanto (1999), who found a negative contribution of the female disadvantage, suggesting that the impact of the femaleness of the establishment on the observed gender wage gap operated mainly through males' wage advantage.

Table 4.2: Male/female log-wage decompositions

Method.	Oaxaca (1973)			Cotton(1988), Neumark (1988)				
	endow.	prices	total	endow.	male adv.	fem. dis.	total	P/gap
1985								
all var.	0.1108	0.1465	0.2574	0.1112	0.0465	0.0997	0.2574	
P	-0.0108	0.0389	0.0281	-0.0028	0.0044	0.0264	0.0281	10.9
1987								
all var.	0.0944	0.1566	0.2510	0.0974	0.0486	0.1049	0.2510	
P	-0.0132	0.0524	0.0391	-0.0021	0.0061	0.0351	0.0391	15.6
1989								
all var.	0.0911	0.1787	0.2698	0.0992	0.0544	0.1162	0.2698	
P	-0.0113	0.0789	0.0675	0.0063	0.0099	0.0513	0.0675	25.0
1991								
all var.	0.0952	0.1942	0.2894	0.1054	0.0617	0.1224	0.2894	
P	-0.017	0.0838	0.0668	0.0027	0.0114	0.0528	0.0668	23.1
1993								
all var.	0.0911	0.1958	0.2869	0.1012	0.0643	0.1214	0.2869	
P	-0.0102	0.0782	0.0681	0.0085	0.0110	0.0485	0.0681	23.7
1995								
all var.	0.1013	0.1630	0.2644	0.1046	0.0612	0.0985	0.2644	
P	-0.0137	0.0397	0.0260	-0.0038	0.0059	0.0240	0.0260	9.9
1997								
all var.	0.0943	0.1615	0.2558	0.0986	0.0619	0.0953	0.2558	
P	-0.0059	0.0479	0.0420	0.0064	0.0073	0.0283	0.0420	16.4
1999								
all var.	0.0944	0.1641	0.2585	0.0990	0.0637	0.0958	0.2585	
P	0.0051	0.0542	0.0593	0.0192	0.0085	0.0316	0.0593	23.0

Notes: P is the proportion of females at the establishment level; *gap* stands for the gender wage gap; 'endow.' is endowments; 'male adv.' and 'fem. dis.' are male advantage and female disadvantage, respectively; 'all var.' is all variables; and, 'Method.' is Methodology.

We have compared the contribution of the variable share of females to the total wage gap with the contribution of industries or occupations. The contribution of occupations taken together or industries taken together to the total gender wage gap is dwarfed by the much larger contribution of the variable share of females.[9]

Table 4.10 in Appendix 4.A reports the decompositions of the wage gap using fixed-effects estimation. Once we account for unobserved heterogeneity across firms, the contribution of the share of females to the total wage gap is, as expected, substantially lower (1 percent to 15 percent contribution when using fixed-effects, instead of 10 percent to

[9]The share of females accounts for 10 percent to 25 percent of the gap, whereas occupations account for -2 percent to 8 percent, and industries account for -8 percent to 0 percent; the only exception is 1985, when the contribution of occupations taken together reaches 19 percent.

Table 4.3: Contributions to the observed gender wage gap (%)

	Oaxaca (1973)				Cotton (1988) and Neumark (1988)					
	all variables		P		all variables			P		
	end.	prices	end.	prices	end.	m. adv.	f. dis.	end.	m. adv.	f. dis.
1985	43.1	56.9	-4.2	15.1	43.2	18.1	38.7	-1.1	1.7	10.3
1987	37.6	62.4	-5.3	20.9	38.8	19.4	41.8	-0.8	2.5	14.0
1989	33.8	66.2	-4.2	29.2	36.8	20.2	43.1	2.3	3.7	19.0
1991	32.9	67.1	-5.9	28.9	36.4	21.3	42.3	0.8	3.9	18.4
1993	31.8	68.2	-3.6	27.3	35.3	22.4	42.3	2.9	3.8	17.0
1995	38.3	61.7	-5.2	15.0	39.6	23.4	37.0	-1.4	2.2	9.1
1997	36.9	63.1	-2.3	18.7	38.5	24.2	37.3	2.5	2.8	11.1
1999	36.5	63.5	2.0	21.0	38.3	24.8	36.8	7.4	3.3	12.2

Notes: P is the proportion of females at the establishment level; 'end.' is endowments; 'm. adv.' and 'f. dis.' are male advantage and female disadvantage, respectively.

25 percent with OLS). However, a remarkable rising trend can still be detected, from a 3 percent contribution in 1985 to 15 percent in 1999. The endowments component still exerts an egalitarian impact on the wage distribution (i.e., a negative contribution to the gender wage gap), showing now a larger magnitude. Therefore, prices continue to be the major force driving the contribution of the variable share of females to the gender wage gap.

In synthesis, for the Portuguese case, segregation remained at stable levels from 1985 to 1999, but nevertheless the degree of femaleness of the establishment tended to become more relevant accounting for wage differences across gender.

The question that would follow is of course what has driven these changes in prices, but at this stage one can only present some speculative reasoning. During the second half of the 1980s the Portuguese economy grew at a very fast pace. A large share of low-paid females in an establishment might have resulted in a larger pie to be distributed among males, in a rent-dissipation argument similar to the one presented by Winter-Ebmer and Zweimüller (1996). Economic growth combined with short supply of qualified labour has indeed led until mid-1990s to rising wage dispersion, with the bottom wages growing slowly, when compared to top wages, which increased very sharply.

An alternative explanation may be derived from Goldin (1990), who analysed specifically the rising female labour force participation and the gender gap in the US. On several fronts, the evidence on Portugal is consistent with Goldin's reasoning. She shows that recent entrants to the labour market tend to be older women, with less labour market experience than the women already in the labour market. The decline in average actual experience would lead to a decline in average females wages and an increase in the gender

gap, particularly if one controls for potential experience and not for actual experience. In Portugal, the average age of employed females indeed increased, from 33.6 years in 1985 to 34.9 in 1999. Also, the gender wage gap, as captured in wage regressions using potential labour market experience, increased up to the early nineties. Progressing in this reasoning, if the women entering the labour market tend to work mainly in establishments that were already employing a high proportion of females, which occurred in Portugal, then segregation would account for an increasing proportion of the wage gap. Furthermore, potential experience contains more error as older women and actually less experienced ones join the labour force, such that the increasing role of segregation would show up as a rising 'price effect'. Evidence on Portugal is also consistent with this piece of the reasoning. In synthesis, rising participation of females from slightly older groups may provide several clues to explain the pattern and trends in the gender pay gap detected in Portugal.

4.6 Conclusion

This chapter analysed gender segregation at the establishment level over fifteen years in Portugal, and its impact on wages and the gender wage gap. A large employer-employee matched data set has been used.

Results point to a high level of systematic gender segregation at the establishment level. A higher proportion of females in the establishment lowers females' wages and, on the contrary, it raises males' wages. The latter outcome contrasts with the evidence available for other countries. Such evidence lends support to wage discrimination type of models. However, comparison of the results obtained using OLS and including firm-specific fixed-effects in the regressions highlights the relevance of sorting of workers into establishments, based on their productivity, to the explanation of the pattern and trend in gender wage setting in Portugal. Similarly, it highlights that discrimination mechanisms are declining over time. The results point to the relevance of taking into account gender segregation of the workforce at the establishment level when analysing the gender wage gap and deciding on policy measures.

Appendix to Chapter 4

4.A Additional tables

Table 4.4: Sample mean values (males)

	1985	1987	1989	1991	1993	1995	1997	1999
Ln hourly wage	5.2596	5.5933	5.8110	6.1595	6.3603	6.4971	6.5818	6.7216
Proportion of females	0.1997	0.2052	0.2145	0.2228	0.2295	0.2363	0.2404	0.2463
Education	5.5132	5.6484	5.9063	6.1435	6.3539	6.6805	7.0088	7.3542
Experience(*)	26.208	26.127	25.507	25.387	25.294	24.810	24.298	24.204
Experience squared(*)	848.27	840.94	810.48	807.32	801.6	777.17	756.13	753.48
Tenure	10.006	10.1162	9.4890	9.2806	9.1149	8.9454	8.5275	8.4567
Tenure squared	178.10	183.91	175.58	175.83	170.14	165.72	158.41	158.44
Ln establishment size	4.6677	4.5950	4.4681	4.3857	4.2385	4.0938	4.0590	4.0027
Lisbon	0.4251	0.4103	0.4007	0.3997	0.3948	0.3805	0.3798	0.3804
Occupations:								
Managers, higher clericals	0.0111	0.0103	0.0113	0.0119	0.0113	0.0311	0.0357	0.0401
Clerical staff	0.0895	0.0883	0.0936	0.0982	0.1007	0.1161	0.1152	0.1269
Commercial staff	0.1357	0.1329	0.1274	0.1248	0.1244	0.1329	0.1248	0.1231
Security and other services	0.0585	0.0588	0.056	0.0568	0.0586	0.0701	0.0715	0.0697
Farmers, agriculture workers	0.0024	0.002	0.0026	0.0023	0.0021	0.0029	0.0035	0.0033
Production workers (group 1)	0.2931	0.2892	0.2861	0.2921	0.2868	0.2933	0.2953	0.2923
Production workers (group 2)	0.1738	0.1718	0.1629	0.1651	0.1603	0.1793	0.1812	0.1759
Production workers (group 3)	0.2118	0.2215	0.2353	0.2201	0.2286	0.1367	0.1363	0.1336
Industries:								
Textiles, clothing, footwear	0.0919	0.0949	0.0938	0.0898	0.0848	0.083	0.078	0.0708
Wood, cork	0.0461	0.046	0.0448	0.0408	0.0407	0.0476	0.0465	0.0441
Paper, print, publishing	0.0272	0.0271	0.0266	0.0263	0.0249	0.0251	0.0238	0.023
Chemical products	0.0480	0.0468	0.0438	0.0368	0.0346	0.0285	0.0255	0.0262
Non-metal minerals	0.0430	0.0406	0.0387	0.0398	0.0382	0.0346	0.0325	0.0336
Primary metals	0.0210	0.0203	0.0176	0.0144	0.013	0.0088	0.0081	0.0077
Machinery, equipment	0.1315	0.1239	0.124	0.1176	0.119	0.1075	0.1125	0.1102
Electricity, gas, water	0.0214	0.0214	0.0157	0.0181	0.0173	0.0168	0.0161	0.0141
Construction	0.1247	0.1257	0.1363	0.1449	0.1488	0.1574	0.1665	0.1634
Wholesale	0.0903	0.0904	0.0893	0.0924	0.0922	0.0865	0.0835	0.0847
Retail	0.0474	0.0491	0.054	0.0563	0.0592	0.0874	0.0884	0.0902
Restaurants, cafes, hotels	0.0309	0.0321	0.0332	0.0336	0.0351	0.0399	0.0389	0.0387
Transportation	0.1083	0.1117	0.1014	0.1096	0.106	0.1002	0.0971	0.0982
Banking, insurance	0.0555	0.0541	0.0601	0.0584	0.0601	0.0595	0.0523	0.0484
Services to firms	0.0176	0.0182	0.0219	0.0261	0.0282	0.0045	0.0048	0.0056
Social, personal services	0.0440	0.0463	0.0484	0.0480	0.0506	0.0675	0.0832	0.1008

Note: (*) Potential experience, computed as $age - education - 6$.

Table 4.5: Sample mean values (females)

	1985	1987	1989	1991	1993	1995	1997	1999
Ln hourly wage	5.0022	5.3423	5.5412	5.8701	6.0735	6.2327	6.3260	6.4631
Proportion of females	0.5639	0.5767	0.5956	0.6082	0.6185	0.6341	0.6455	0.6505
Education	5.4763	5.7060	6.0359	6.3121	6.5439	7.0174	7.3776	7.7936
Experience(*)	22.168	22.272	21.626	21.267	21.240	21.278	21.125	21.118
Experience squared(*)	627.35	631.66	603.7	588.88	588.39	592.58	592.64	597.43
Tenure	8.9576	9.0402	8.3458	7.8066	7.5880	7.7406	7.4244	7.2951
Tenure squared	136.07	143.12	135.83	129.95	124.24	127.21	123.15	122.81
Ln establishment size	4.6199	4.5449	4.4241	4.3596	4.2646	4.1835	4.1423	4.1082
Lisbon	0.4018	0.3948	0.3790	0.3733	0.3709	0.3621	0.3588	0.3702
Occupations:								
Managers, higher clericals	0.0127	0.0139	0.0166	0.0171	0.0202	0.0305	0.0351	0.0405
Clerical staff	0.0677	0.0704	0.0742	0.0793	0.0814	0.0718	0.0721	0.0814
Commercial staff	0.2124	0.2019	0.1968	0.1894	0.1852	0.2145	0.2067	0.2125
Security and other services	0.0768	0.0817	0.0825	0.0861	0.0967	0.1400	0.1624	0.1713
Farmers, agriculture workers	0.0016	0.0012	0.0023	0.0018	0.0014	0.0016	0.0021	0.0018
Production workers (group 1)	0.2469	0.2506	0.2541	0.2649	0.2512	0.2427	0.2388	0.2219
Production workers (group 2)	0.1249	0.1143	0.0954	0.0812	0.0707	0.0943	0.0797	0.0697
Production workers (group 3)	0.2507	0.2575	0.2681	0.2685	0.2824	0.1886	0.1871	0.1848
Industries:								
Textiles, clothing, footwear	0.3288	0.3287	0.3279	0.3210	0.3017	0.2858	0.2633	0.2366
Wood, cork	0.027	0.0248	0.0234	0.0224	0.0222	0.0261	0.0258	0.0252
Paper, print, publishing	0.0216	0.0217	0.02	0.018	0.0172	0.0178	0.0159	0.0157
Chemical products	0.0396	0.0363	0.0318	0.0275	0.0248	0.0198	0.0167	0.0177
Non-metal minerals	0.0248	0.0233	0.0222	0.0245	0.024	0.0238	0.0223	0.0226
Primary metals	0.0049	0.0041	0.0035	0.0028	0.0024	0.0016	0.0014	0.0016
Machinery, equipment	0.0689	0.0657	0.0600	0.0633	0.0627	0.0674	0.0684	0.0716
Electricity, gas, water	0.0068	0.0069	0.0045	0.0050	0.0046	0.0042	0.0043	0.0037
Construction	0.0129	0.0117	0.0141	0.0161	0.0162	0.0157	0.0167	0.0171
Wholesale	0.0687	0.0654	0.0643	0.0648	0.0647	0.0594	0.0579	0.0587
Retail	0.0627	0.0651	0.0679	0.0716	0.0785	0.0948	0.1043	0.1118
Restaurants, cafes, hotels	0.0534	0.0542	0.0563	0.0574	0.063	0.0715	0.0710	0.0684
Transportation	0.0585	0.0582	0.0534	0.0524	0.0489	0.0433	0.0407	0.0388
Banking, insurance	0.0421	0.0402	0.0447	0.0419	0.0436	0.0440	0.0396	0.0398
Services to firms	0.0139	0.0162	0.0210	0.0256	0.0284	0.0046	0.0048	0.0056
Social, personal services	0.0916	0.1112	0.1238	0.1275	0.1441	0.1683	0.1984	0.2217

Note: (*) Potential experience, computed as $age - education - 6$.

Table 4.6: Ordinary least squares regressions (males)

	1985		1987		1989		1991	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	0.0296	13.7	0.0356	16.1	0.0297	13.2	0.0441	18.1
Education	0.0506	290.6	0.0552	305.2	0.0623	331.8	0.0656	328.1
Experience	0.0268	197.6	0.0287	204.3	0.0296	207.3	0.0281	182.3
Experience sq./100	-0.0353	-162.0	-0.0374	-165.0	-0.0383	-164.0	-0.0362	-142.6
Tenure/10	0.1036	86.4	0.1047	85.5	0.1033	81.0	0.0982	68.8
Tenure squared/100	-0.0129	-36.3	-0.0122	-33.3	-0.0118	-30.0	-0.0103	-23.3
Ln establishment size	0.0575	245.2	0.0604	250.2	0.0551	216.9	0.0543	188.4
Region: Lisbon	0.0726	98.2	0.0686	89.4	0.0763	93.7	0.0985	109.9
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	862137		860395		889362		885135	
Adjusted R^2	0.6234		0.6373		0.5995		0.5596	
	1993		1995		1997		1999	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	0.0262	10.4	0.0344	14.7	0.0147	6.6	-0.0127	-6.4
Education	0.0682	331.8	0.0474	216.9	0.0456	223.3	0.0459	243.8
Experience	0.0279	172.7	0.0272	176.1	0.0269	182.8	0.0254	193.0
Experience sq./100	-0.0350	-130.6	-0.0363	-141.1	-0.0357	-143.7	-0.0336	-151.5
Tenure/10	0.1015	64.6	0.1052	67.9	0.1304	89.0	0.1420	105.2
Tenure squared/100	-0.0134	-27.3	-0.0125	-25.8	-0.0188	-41.2	-0.0203	-48.1
Ln establishment size	0.0554	185.4	0.0570	191.8	0.0584	212.7	0.0569	215.4
Region: Lisbon	0.1104	116.3	0.1037	112.3	0.0923	102.8	0.0934	112.8
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	868326		859522		923256		947059	
Adjusted R^2	0.5444		0.6041		0.5951		0.6343	

Note: In each regression the dependent variable is log hourly wage.

Table 4.7: Ordinary least squares regressions (females)

	1985		1987		1989		1991	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	-0.0393	-17.6	-0.0552	-24.3	-0.1027	-45.1	-0.0936	-38.6
Education	0.0475	178.1	0.0564	203.0	0.0610	229.3	0.0662	235.6
Experience	0.0155	81.8	0.0166	87.7	0.0175	99.5	0.0161	89.4
Experience sq./100	-0.0192	-59.6	-0.0199	-61.7	-0.0202	-65.2	-0.0180	-54.8
Tenure/10	0.1209	72.0	0.1213	72.8	0.1186	71.3	0.1169	64.8
Tenure squared/100	-0.0216	-38.8	-0.0199	-35.5	-0.0195	-32.0	-0.0173	-26.2
Ln establishment size	0.0458	138.6	0.0486	147.3	0.0444	133.3	0.0443	126.4
Region: Lisbon	0.0764	70.4	0.0679	62.3	0.0660	61.0	0.0795	69.4
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	402523		424697		477440		507748	
Adjusted R^2	0.6594		0.6498		0.6355		0.5864	
	1993		1995		1997		1999	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	-0.1003	-38.7	-0.0283	-12.0	-0.0596	-26.8	-0.0960	-50.7
Education	0.0710	246.4	0.0434	170.1	0.0411	177.0	0.0423	213.5
Experience	0.0166	87.2	0.0176	98.3	0.0183	109.6	0.0169	126.2
Experience sq./100	-0.0182	-52.2	-0.0228	-68.3	-0.0243	-78.6	-0.0215	-88.8
Tenure/10	0.1297	63.3	0.1284	67.8	0.1368	79.6	0.1482	105.2
Tenure squared/100	-0.0228	-30.9	-0.0221	-33.5	-0.0236	-39.9	-0.0254	-52.7
Ln establishment size	0.0473	135.3	0.0548	164.7	0.0532	174.7	0.0440	172.4
Region: Lisbon	0.0927	75.8	0.0769	68.4	0.0715	66.8	0.0637	72.5
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	524732		562909		634009		675553	
Adjusted R^2	0.5631		0.6317		0.6118		0.6833	

Note: In each regression the dependent variable is log hourly wage.

Table 4.8: Firm fixed-effects regressions (males)

	1985		1987		1989		1991	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	0.1169	24.2	0.0771	15.5	0.0839	15.7	0.1065	18.3
Education	0.0397	248.6	0.0422	255.9	0.0478	272.4	0.0499	265.8
Experience	0.0229	205.6	0.0243	205.1	0.0264	212.1	0.0257	189.1
Experience sq./100	-0.0300	-176.0	-0.0300	-170.4	-0.0300	-173.3	-0.0300	-152.1
Tenure/10	0.0920	85.4	0.0880	78.5	0.0850	70.1	0.0810	58.8
Tenure squared/100	-0.0100	-26.3	-0.0100	-21.2	-0.0100	-14.5	0.0000	-10.1
Ln establishment size	-0.0023	-6.1	0.0012	3.0	-0.0009	-2.0	0.0014	2.9
Region: Lisbon	0.0129	11.8	0.0183	16.1	0.0274	21.4	0.0304	21.3
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	862137		860395		889362		885135	
Adjusted R^2	0.800		0.800		0.770		0.730	
	1993		1995		1997		1999	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	0.0633	10.1	0.0299	5.1	-0.0032	-0.6	0.0095	1.9
Education	0.0523	266.1	0.0350	177.7	0.0342	185.5	0.0350	207.2
Experience	0.0261	181.3	0.0244	177.6	0.0247	191.1	0.0237	208.1
Experience sq./100	-0.0300	-141.0	-0.0300	-143.1	-0.0300	-151.4	-0.0300	-165.1
Tenure/10	0.0850	56.1	0.0930	61.4	0.1120	78.5	0.1170	90.5
Tenure squared/100	-0.0100	-13.4	-0.0100	-20.1	-0.0100	-32.8	-0.0100	-37.8
Ln establishment size	0.0038	7.2	-0.0005	-1.0	0.0021	4.2	0.0048	9.1
Region: Lisbon	0.0339	22.3	0.0417	26.7	0.0383	26.5	0.0350	25.2
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	868326		859522		923256		947059	
Adjusted R^2	0.710		0.760		0.760		0.790	

Note: In each regression the dependent variable is log hourly wage.

Table 4.9: Firm fixed-effects regressions (females)

	1985		1987		1989		1991	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	0.0290	5.1	0.0232	3.8	0.0171	2.6	0.0153	2.2
Education	0.0346	146.9	0.0401	170.7	0.0454	188.3	0.0499	195.6
Experience	0.0116	73.8	0.0128	80.1	0.0143	92.8	0.0136	84.2
Experience sq./100	-0.0100	-56.7	-0.0200	-59.9	-0.0200	-65.1	-0.0200	-54.7
Tenure/10	0.1080	65.1	0.1110	67.1	0.1060	63.6	0.1100	60.6
Tenure squared/100	-0.0200	-32.1	-0.0200	-30.8	-0.0100	-24.5	-0.0100	-23.2
Ln establishment size	0.0040	5.9	0.0090	12.6	0.0048	7.3	0.0121	16.7
Region: Lisbon	0.0241	11.5	0.0277	13.5	0.0254	11.2	0.0335	15.1
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	402523		424697		477440		507480	
Adjusted R^2	0.823		0.810		0.790		0.740	
	1993		1995		1997		1999	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
Proportion of females	-0.0058	-0.8	-0.0275	-3.9	-0.0421	-6.7	-0.0545	-10.0
Education	0.0545	206.5	0.0316	132.5	0.0304	138.4	0.0314	171.3
Experience	0.0143	82.4	0.0145	91.2	0.0151	101.5	0.0146	124.6
Experience sq./100	-0.0200	-52.3	-0.0200	-66.1	-0.0200	-75.6	-0.0200	-90.1
Tenure/10	0.1230	59.3	0.1270	66.1	0.1360	76.5	0.1360	96.2
Tenure squared/100	-0.0200	-27.5	-0.0200	-32.0	-0.0200	-38.4	-0.0200	-48.0
Ln establishment size	0.0088	11.4	-0.0011	-1.6	0.0013	1.9	0.0055	8.9
Region: Lisbon	0.0375	16.3	0.0358	15.9	0.0481	23.9	0.0310	19.0
Occupation(9 categories)	yes		yes		yes		yes	
Industry (17 categories)	yes		yes		yes		yes	
Observations	524732		562909		634009		675553	
Adjusted R^2	0.710		0.770		0.750		0.820	

Note: In each regression the dependent variable is log hourly wage.

Table 4.10: Contribution of the share of females to the total gender wage gap, under alternative estimation methods

Oaxaca or Cotton-Neumark decomposition (total component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	10.9	15.6	25.0	23.1	23.7	9.9	16.4	23.0
fixed effects	2.7	1.0	2.9	5.0	6.3	9.3	10.3	14.6

Oaxaca decomposition (endowments component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	-4.2	-5.3	-4.2	-5.9	-3.6	-5.2	-2.3	2.0
fixed effects	-16.5	-11.4	-11.9	-14.2	-8.6	-4.5	0.5	-1.5

Oaxaca decomposition (price component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	15.1	20.9	29.2	28.9	27.3	15.0	18.7	21.0
fixed effects	19.2	12.4	14.7	19.2	14.9	13.8	9.8	16.1

Cotton-Neumark decomposition (endowments component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	-1.1	-0.8	2.3	0.8	2.9	-1.4	2.5	7.4
fixed effects	-12.6	-8.8	-8.6	-9.8	-5.1	-1.1	3.0	2.7

Cotton-Neumark decomposition (male advantage component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	1.7	2.5	3.7	3.9	3.8	2.2	2.8	3.3
fixed effects	2.2	1.5	1.9	2.6	2.1	2.0	1.5	2.5

Cotton-Neumark decomposition (female disadvantage component)

	1985	1987	1989	1991	1993	1995	1997	1999
OLS	10.3	14	19	18.4	17	9.1	11.1	12.2
fixed effects	13.1	8.3	9.6	12.2	9.3	8.3	5.8	9.4

Chapter 5

Micro foundations for wage flexibility: wage insurance at the firm level

5.1 Introduction

The impact of high wage flexibility reducing economic fluctuations and improving macroeconomic performance has been highlighted in the literature, where wage flexibility has invariably been evaluated as the responsiveness of wages to aggregate conditions, namely the unemployment rate. However, in the terminology of Faggio and Nickell (2005), wage flexibility has two different aspects: the responsiveness of wages to labour market conditions, and the responsiveness of wages within a firm to idiosyncratic shocks to its productivity or its output. They concentrate on the first aspect. The second aspect can be understood as the micro foundations for wage flexibility, the issue under analysis in the current study. We focus on Portugal, pointed out as one of the OECD economies with highest wage flexibility, despite its labour market regulations.

More precisely, we will provide an answer to the questions: What is the responsiveness of wages to shocks to firm output? I.e., to which extent do firms provide wage insurance to their workers, insulating them from fluctuations in product markets? Which firm and worker attributes are associated with a higher degree of wage flexibility at the firm level? A very precise hypothesis has been stated by Faggio and Nickell (2005): national collective bargaining is associated with lower responsiveness of wages to labour market conditions. We will check whether workers covered by national bargaining agreements also see their wages react less to firm level idiosyncratic shocks than workers covered by more decentralized agreements. At first sight that might be expected, but it is not necessarily the case. Indeed, Teulings (1997) has argued that in a corporatist setting for

wage bargaining, firms can delegate on trade unions the task of adjusting contracts to macro level shocks, while then promoting adjustments to firm idiosyncratic shocks. Also, Cardoso and Portugal (2005) have shown that firms are able to overcome the constraints imposed by collective bargaining by adjusting the actual wage paid on top of the bargained wage. We will also inspect the impact of other labour market regulations, namely the minimum wage, on wage flexibility. Other hypotheses can be derived from the wage insurance literature, which has shown that the share of risk borne by the firm and the worker depends on factors such as: the persistence of the shocks hitting the firm; workers' and firms' preferences, namely their degree of risk-aversion; the sensitivity of firm output to worker effort; the likelihood of bankruptcy.

An empirical test on such theories depends crucially on the quality and detail of the data available. We use a longitudinal matched employer-employee data set of remarkable quality, which matches all the firms and workers in the manufacturing and services private sectors. Given its nature, problems commonly faced by longitudinal data sets, such as panel attrition and under- or over-representation of certain groups, are avoided. Also, the legal requirement for the data to be posted in a visible location within the company contributes to its reliability, reducing measurement errors.

Guiso et al. (2005) have devised an empirical strategy to quantify the impact of temporary and permanent firm-level shocks on wages, which relies on longitudinal matched employer-employee data to estimate dynamic panel data models. We will follow their strategy to quantify the wage response to firm-level permanent and transitory shocks. We will then explore the forces that shape wage flexibility at the firm level, in particular the role of the institutional setting.

After the brief revision of the literature that follows, Section 5.3 describes the institutional framework for wage setting in Portugal and Section 5.4 describes the data. Sections 5.5 to 5.8 summarize the empirical model and present the results, before concluding comments are presented in Section 5.9.

5.2 Wage insurance in the previous literature

Insurance models can explain why wages do not adjust as much as predicted by spot market theory, after changes in the demand for the firm output. The underlying idea is that firms, being risk neutral, commit to paying a pre-defined wage to their risk averse workers, independently of product market fluctuations. Such strategy is profit maximizing because risk-averse workers will accept a non-stochastic wage lower than the expected value of a stochastic wage. Early models have been developed by Baily (1974), Gordon

(1974), and Azariadis (1975). Other models predict relatively smaller insurance provision. Gamber (1988) allows for firm bankruptcy, which constrains the capacity of the firm to provide insurance to the workers, and distinguishes between temporary and permanent shocks, in a two-period model. In his model, real wages react more to permanent shocks than to temporary ones. In case of temporary shocks, the firm wishing to smooth the wage of the worker over time can promote a relatively small wage adjustment in the period the shock occurs, deferring the rest of the adjustment to the following period.

A central issue that follows concerns the enforceability of insurance contracts. For example, if worker performance is not verifiable, the firm may gain from declaring that it is below its actual level and reneging the contract, thus paying a wage lower than the insurance wage. Similarly, if worker mobility is allowed, the worker might gain from reneging the contract and accepting a better outside offer. Implicit contract theory has established conditions under which it is in the firm's and in the worker's interest to stick to the contract. Basically, workers and firms will respect the contract as long as its long-run gains outweigh the short term benefit from reneging it. The insurance wage could therefore fluctuate between the level strictly required to prevent the firm from dismissing the worker and, by a similar reasoning, the level strictly required to prevent the worker from quitting. The latter case holds when contracts are not binding on the worker, whereas the former holds when contracts are not binding on the firm.[1]

Empirical studies relied initially on aggregate industry data (Gamber, 1988; Christofides and Oswald, 1992; Blanchflower et al., 1996), progressing to use firm-level averages (Hildreth and Oswald, 1997; Nickell and Wadhwani, 1990). Beaudry and DiNardo (1991) use individual worker data, but their indicator of market conditions is computed at the aggregate or industry level. Similarly, Weinberg (2001) uses individual data, but relies on a measure of shocks defined at the industry level, to analyse wage and employment fluctuations at the industry level in response to demand shocks. Devereux (2005) relies on panel data on workers to quantify the impact of industry-level demand shocks on wages, finding that industry wages respond positively to changes in industry employment. Faggio and Nickell (2005) use worker longitudinal data to quantify the impact of changes in labour market conditions at the regional level on wages. Finally, Guiso et al. (2005) have set a new benchmark in the analysis of this issue. The ingenious empirical identification strategy followed relies on longitudinal matched employer-employee data to estimate dy-

[1]For an early overview of contract theory, see Rosen (1985). Weiss (1984) has considered the role of mobility costs preventing workers from quitting and thus enabling firms to provide wage insurance; Holmstrom (1981) and Thomas and Worrall (1988) model the consequences of the loss of reputation by firms that renege on a contract; in the model by Harris and Holmstrom (1982), firms learn about worker ability and adjust the wage to the outside market to prevent the worker from quitting.

namic panel data models and quantify the impact of temporary and permanent firm-level shocks on wages. They found that firms provide full insurance against temporary shocks, while providing only partial insurance against permanent shocks.[2]

We will follow Guiso et al. (2005) and the basic intuition behind the procedure implemented in sections 5.5 to 5.8 is the following. First, consistent estimates of the residuals from a wage and a firm's performance regression are retrieved. Second, we regress idiosyncratic shocks to wages on idiosyncratic shocks to firm's performance, and evaluate the level of insurance provided by the firm to both temporary and permanent shocks. Finally, we check for heterogeneity of the wage reaction according to firm and worker attributes.

5.3 Wage setting institutions in Portugal

The Portuguese labour market is characterized by a high level of employment rigidity and high wage flexibility. In fact, its strict job protection legislation, covering issues such as advance notice required before dismissal, severance pay, and the rules on use of fixed-term or temporary contracts, invariably place the country among the OECD economies with highest employment rigidity (see, for example, OECD, 1999). On the contrary, the country ranks among the OECD economies with highest wage flexibility (see OECD, 1992), since wages are highly responsive to the unemployment rate, despite the regulated framework.

Even though union membership has declined, from 61 percent in 1970 and 1980 to 32 percent in 1990 (OECD, 1994b, p.184), collective bargaining covers almost all of the workforce. This wide coverage results from widespread mechanisms of extension of contracts: most often, employers who subscribe to an agreement apply it to all of their workforce, irrespective of the worker union membership status; employers or workers representatives can join an existing agreement, subscribing to a contract they had initially not signed; moreover, the Government can impose mandatory extensions of existing contracts, when workers are not covered by a trade union, when one of the parties refuses to negotiate or negotiation is obstructed in any other way.

Studies at the micro level have identified sources of wage flexibility under this regulated

[2]In a competitive wage setting where wages are perfectly determined by the market, full insurance resembles perfect market flexibility where wages do not depend on firms' idiosyncracies. In such context, the term "full insurance" can be misleading. In a situation where wages have some degree of dependence on firms' specific performance, what is defined as partial insurance in the current terminology might as well be a rent seeking behavior of incumbents who are able to negociate wages on top of the market wage. However, partial insurance translates the idea since workers are trading a full insured non-stochastic wage by potential gains associated with some wage flexibility. As such, we can call it partial insurance even though it is the result of the bargaining between firms and their workers.

setting. In particular, Cardoso and Portugal (2005) have found that wages set by collective bargaining reflect to a high extent the degree of power of the partners negotiating, but subsequent firm-specific arrangements reduce the returns to union power, adjusting wages to the conditions that prevail at the micro level. Also, Cardoso (1999) had found that the returns to different worker attributes vary widely across firms.

As a rule, wage negotiations are held yearly and the wage updates take effect in January each year.[3]

5.4 Data set

Quadros de Pessoal is a matched employer-employee data set gathered by the Ministry of Employment, based on an inquiry that every company with wage-earners is legally obliged to fill in. Public administration and domestic service are not covered, and the coverage of agriculture is low, given its low share of wage-earners. For the remaining sectors, the mandatory nature of the survey leads to an extremely high response rate. Each year, around two million workers and 100 to 200 thousand companies are covered. Data for 1991 to 2000 are used.

Reported data cover the firm and all the workers engaged in the firm in a reference week (whether wage-earner, unpaid family member of owner working in the firm). Reported variables include the firm's location, industry, employment, sales volume, ownership structure, and date of creation, and the worker's gender, schooling, age, occupation, seniority, several components of wage, duration of work, and collective bargaining contract.

A worker identification code, based on a transformation of the social security number, enables tracking him/her over time. Extensive checks have been performed to guarantee the accuracy of the data, using gender, date of birth, and highest schooling level achieved. A firm identification code enables tracking it over time. Based in particular on the location of the firm and its official identification codes, extensive controls are implemented by the data gathering agency to guarantee that a firm is not assigned a different number later on.

Details on the construction of the database, sample sizes, and descriptive statistics are presented in Appendices 5.A, 5.B and 5.C.

[3]Further details on the Portuguese labour market can be found in Section 4.3.

5.5 Firm performance

Based on the specification used by Guiso et al. (2005), firms' performance is modeled as

$$sales_{jt} = \gamma_t + \rho sales_{j,t-1} + X'_{jt}\Gamma + \eta_j + \epsilon_{jt} \tag{5.1}$$

where $sales_{jt}$ is the logarithm of sales of firm j in period t, X_{jt} is a vector of firm characteristics that includes a set of industry and location dummies, γ_t represents period t specific constant, ρ and Γ are parameters to be estimated, η_j is the firm specific effect, and ϵ_{jt} is the remaining component of the error term.

A major issue concerns the empirical measurement of fluctuations in product markets. The shock affecting the firm has been defined using: the industry output price (Gamber, 1988; Christofides and Oswald, 1992); the industry profit (Blanchflower et al., 1996; Christofides and Oswald, 1992); firm profits, in studies that rely however on wage data also aggregated for the firm level (Hildreth and Oswald, 1997; Nickell and Wadhwani, 1990). Abowd and Lemieux (1993) rely on a set of assumptions to compute a profitability variable (quasi-rents per worker) at the firm level, and use the price of exports and imports at the industry level to instrument it. Guiso et al. (2005) use value added instead of profits, arguing that it is the variable directly subject to stochastic fluctuations, being more reliable than profits. A similar option was taken by Estevão and Tevlin (2003), who nevertheless used industry data. Holzer and Montgomery (1993) used firm sales, with wages averaged for the firm level. We use sales as our indicator of firm performance, arguing that it captures demand uncertainty, as shocks in product demand are directly reflected in changes in sales. Given fluctuations in demand, output could remain unchanged if prices would adjust fully and instantaneously, but since that is not the case, output will undergo fluctuations (Baily, 1974). Sales were deflated using the GDP deflator.

Estimation of equation (5.1) by OLS or the usual panel models, fixed or random effects, is inconsistent in the presence of the lagged dependent variable, since, by definition, $sales_{j,t-1}$ is correlated with η_j. We follow Arellano and Bond (1991), taking first differences to eliminate the fixed effect, and then estimating equation (5.1) using a generalized method of moments (GMM) procedure. The set of instruments include $sales_{j,t-3}$ and earlier levels of this variable. The remaining regressors are treated as exogenous, and introduced in levels as instruments. The results for the 1–step GMM estimation procedure are reported in Table 5.1.

The use of this method calls for some discussion. This solution has poor finite sample properties concerning bias and precision when the available instruments are weak. Blundell and Bond (1998) show that the solution of Arellano and Bond (1991) has a large

Table 5.1: Sales regression

Variable	Estimate	
Log sales at $t-1$	.473	(.022)
Region dummies	8.534	[.074]
Industry dummies	72.54	[.000]
Year dummies	151.3	[.000]
Sargan	37.1	[.093]
Sargan-df	27	
AR(1)	-21.41	[.000]
AR(2)	5.244	[.000]
AR(3)	.718	[.473]
AR(4)	-.833	[.405]
AR(5)	-.153	[.879]
AR(6)	.665	[.506]
AR(7)	-.672	[.501]
Observations	94365	
Firms	17097	

Notes: The regression has been estimated by the first-differenced GMM procedure discussed in Arellano and Bond (1991). The instruments are discussed in the text. The dependent variable is log real sales. Robust standard errors reported in parentheses; p-values in brackets. For region, industry and year dummies, the joint $F-statistic$ is reported. Sargan-df stands for the degrees of freedom of the Sargan test. AR(n) shows the nth order test for serial correlation in the first-differenced residuals.

Table 5.2: Firms' autocovariances

τ	$E(\Delta\epsilon_{jt}, \Delta\epsilon_{j,t-\tau})$	Standard error
0	.7795	.0151
1	-.3096	.0080
2	-.0653	.0103
3	.0031	.0076
4	.0083	.0073
5	-.0051	.0070
6	-.0020	.0067
7	-.0009	.0067

Note: The autocovariances are computed using all years pooled.

downward bias when the time series are persistent and the number of periods is small, and argue for the implementation of a system GMM estimation, for first-differences and levels. In our case, this solution is not feasible given the structure of the error component ϵ_{jt} assumed later on.[4]

The persistence of sales over time is represented by a coefficient on lagged sales of 0.47. Our results indicate that industry dummies are jointly significant, just like time dummies and region dummies. According to the Sargan test, we do not reject the validity of our instruments at the 1% and 5% levels. The serial correlation in the first-differenced residuals indicates that we should be using lagged levels of sales dated $t-3$ and earlier, as we do.

In Table 5.2 we report the autocovariance structure for $\Delta\epsilon_{jt}$. The results confirm our choice of instruments. After 3 lags the covariance of first-differenced residuals is insignificant. These results are of particular interest for the specification of the structure of the error term which will take place in Section 5.7.

5.6 Worker earnings

Workers' wages are specified as

$$wage_{ijt} = K_{ijt}^{'}\Phi + \varphi_i + \alpha P_{jt} + \beta T_{jt} + \psi_{ijt} \tag{5.2}$$

[4]In Section 5.7 we define $\epsilon_{jt} = \zeta_{jt} + \tilde{\nu}_{jt} - \theta\tilde{\nu}_{j,t-1}$ and $\zeta_{jt} = \zeta_{j,t-1} + \tilde{u}_{jt}$, which implies that $Cov(\epsilon_{jt}, \Delta\epsilon_{j,t-\tau}) \neq 0$. This renders infeasible the implementation of the system GMM estimation.

Table 5.3: Wage regression

Variable	Estimate	
Log wage at $t-1$	.692	(0.083)
Region dummies	13.11	[.108]
Industry dummies	24.54	[.220]
Year dummies	126.5	[.000]
Sargan	27.53	[.121]
Sargan-df	20	
AR(1)	-10.85	[.000]
AR(2)	5.180	[.000]
AR(3)	-1.837	[.066]
AR(4)	1.745	[.081]
AR(5)	-.652	[.515]
AR(6)	.607	[.544]
AR(7)	-1.213	[.2225]
Observations	98655	
Individuals	30657	

Notes: The dependent variable is log real monthly wage. See the note to Table 5.1.

where $wage_{ijt}$ stands for the logarithm of monthly wage of worker i engaged in firm j in period t, and K includes industry, region and year dummies, as well as age and age squared. The first component of the error term is the worker specific effect, φ_i. Following Guiso et al. (2005), we include in the wage regression the permanent and transitory components of firm specific shock, P_{jt} and T_{jt}, respectively. The parameters α and β capture the impact of these shocks on wages. Finally, ψ_{ijt} is the remaining component of the error term not captured by the worker specific effect or the firm specificities.

To replicate Guiso et al.'s (2005) strategy to identify α and β we need to multiply equation (5.2) by $(1-\rho L)$, where L is the lag operator. The transformed wage equation is defined as

$$wage_{ijt} = \rho wage_{ij,t-1} + (1-\rho L)K'_{ijt}\Phi + (1-\rho L)(\varphi_i + \alpha P_{jt} + \beta T_{jt} + \psi_{ijt}) \tag{5.3}$$

The direct implication is that we introduce state dependence on wages in the equation to be estimated. The presence of the lagged dependent variable on the right hand side as a result of this transformation brings about an endogeneity problem. In order to solve this issue, and as in the case of equation (5.1), we use Arellano and Bond first-differenced GMM procedure to obtain consistent estimates.

The transformation applied to equation (5.2) leads to the composite error term $\omega_{ijt} =$

Table 5.4: Workers' autocovariances

τ	$E(\Delta\omega_{jt}, \Delta\omega_{j,t-\tau})$	Standard error
0	.0536	.0012
1	-.0253	.0008
2	-.0034	.0009
3	-.0009	.0007
4	.0005	.0006
5	-.0001	.0008
6	.0010	.0010
7	-.0017	.0014

Notes: The autocovariances are computed using all years pooled. $\Delta\omega_{ijt}$ is the first-differenced composite residual from equation (5.2).

$(1-\rho L)(\varphi_i + \alpha P_{jt} + \beta T_{jt} + \psi_{ijt})$. It follows from the above setup that it includes worker's idiosyncrasy, together with the permanent and transitory shocks to firm's performance. The transformation $(1-\rho L)$ imposes a structure on wage residuals that is related to the covariance structure present in equation (5.1). The resulting components and specific covariance structure will be the basis for the identification of α and β, and will be discussed in detail in section 5.7. For the moment, we concentrate on the estimation and analysis of the first-differenced composite error term $\Delta\omega_{ijt}$.

We use levels of wage lagged 4 periods and earlier as instruments for first-differenced equations. The remaining regressors are assumed exogenous and introduced in levels. The results for the 1–step first-differenced GMM estimation are reported in Table 5.3.[5]

The coefficient on lagged wage is 0.69, indicating higher persistence than for sales. Industry dummies are not jointly significant, while region dummies are marginally insignificant at the 10% level. The test for overidentifying restrictions does not reject our instruments. Table 5.4 reports the covariance structure of first-differenced residuals associated with equation (5.2), $\Delta\omega_{ijt}$. First-differencing implies that $\Delta\omega_{ijt}$ lacks φ_i; i.e., it is defined only as a function of the remaining three components of the error term in equation (5.2). The results support our choice of instruments in Table 5.3.

[5]We have considered each employment spell as a pair worker-firm, since we are interested in the provision of wage insurance by a given firm, and not the overall insurance the worker may enjoy when switching firms.

5.7 Insurance provision by the firm

To quantify the insurance provided by firms to their workers we need first to estimate the sensitivity parameters, α and β, and then to estimate the different variance components of the error terms associated with equations (5.1) and (5.2). Throughout the section, we borrow the formulation and estimation strategy proposed by Guiso et al. (2005), adjusting for the specificities of our analysis. The main findings are reported in Table 5.5.

We start by showing in Panel A the covariance structures in the matched sample of firms and workers, which contains 71585 observations. The first two columns report results similar to those shown in Tables 5.2 and 5.4. The last column shows that the covariance between the worker's and the firm's lagged shocks is positive and significant, which is a first indication that firms do not provide full insurance to their workers.

We proceed with the estimation of α and β and report the results in Panel B. To assess insurance within the firm we now focus our attention on the relation between changes in workers' residuals, $\Delta\omega_{ijt}$, and changes in the firms' residuals, $\Delta\epsilon_{jt}$. Firms' error term, ϵ_{jt}, is formulated as the sum of two components: a random walk and a MA(1), such that $\epsilon_{jt} = \zeta_{jt} + \tilde{\nu}_{jt} - \theta\tilde{\nu}_{j,t-1}$, where $\zeta_{jt} = \zeta_{j,t-1} + \tilde{u}_{jt}$. By assuming that $E(\tilde{u}_{jt}^2) = \sigma_{\tilde{u}}^2$, $E(\tilde{\nu}_{jt}^2) = \sigma_{\tilde{\nu}}^2$ for all t, $E(\tilde{\nu}_{js}\tilde{\nu}_{jt}) = E(\tilde{u}_{js}\tilde{u}_{jt}) = 0$ for $s \neq t$, and $E(\tilde{\nu}_{js}\tilde{u}_{jt}) = 0$ for all s and t, we expect that after two periods the autocovariance of $\Delta\epsilon_{jt}$ goes to zero. Empirically, Table 5.2 gives support to this specification, since we observe that autocovariances are zero for lags above 2, and non-zero for two or less lags.

The worker specific effect φ_i drops out of $\Delta\omega_{ijt}$. From Guiso et al., first-differences of permanent and transitory shocks to firms performance are defined as $\Delta P_{jt} = (1-\rho)^{-1}\tilde{u}_{jt}$ and $\Delta T_{jt} = (1-\rho L)^{-1}[(1-\theta L)\Delta\tilde{\nu}_{jt} - (1-\rho)^{-1}\rho\Delta\tilde{u}_{jt}]$, respectively. Finally, the last component of the error term in equation (5.2) is defined as $\psi_{ijt} = \vartheta_{ijt} + \xi_{ijt} - \lambda\xi_{ij,t-1}$, with $\vartheta_{ijt} = \vartheta_{ij,t-1} + \mu_{ijt}$. This specification is also not rejected by the results for the autocovariances in $\Delta\omega_{ijt}$, Table 5.4.

At the core of the estimation strategy lies two instrumental variables regressions, whose specific instruments allow for the identification of the parameters of interest; i.e., α, the sensitivity of wages to permanent shocks, and β, the sensitivity of wages to transitory shocks. In both cases, the dependent variable is $\Delta\omega_{ijt}$, and the explanatory variable is $\Delta\epsilon_{jt}$. Consistent estimates of these variables are obtained from sales and wage regressions presented in Tables 5.1 and 5.3, respectively. Under the defined covariance structure for $\Delta\omega_{ijt}$ and $\Delta\epsilon_{jt}$, Guiso et al. (2005) show that $(\sum_{\tau=-2}^{2}\Delta\epsilon_{j,t+\tau})^k$ is a valid set of instruments for the first regression where we estimate α, while in the second regression the estimation of β can be based on the instruments $(\Delta\epsilon_{j,t+1})^k$.

To estimate both α and β we have used the feasible efficient GMM procedure, con-

Table 5.5: Testing for insurance

A. Covariances			
τ	$E(\Delta\omega_{jt}, \Delta\omega_{j,t-\tau})$	$E(\Delta\epsilon_{jt}, \Delta\epsilon_{j,t-\tau})$	$E(\Delta\omega_{jt}, \Delta\epsilon_{j,t-\tau})$
0	.0545	.7174	-.0012
	(.0014)	(.0265)	(.0010)
1	-.0256	-.2912	.0035
	(.0009)	(.0143)	(.0010)
B. Sensitivity to permanent and transitory shocks			
	Permanent shock	Transitory shock	
Sensitivity	.0924	-.0011	
	(.0446)	(.0019)	
Observations	25667	55077	
J-test	[.5405]	[.1919]	
F-test	[.0019]	[.0000]	
Exogeneity test	[.0422]		
C. Variance components and insurance coverage			
	Firm		Worker
$\sigma^2_{\tilde{u}}$	.1325	σ^2_{μ}	.0058
	(.0203)		(.0113)
$\sigma^2_{\tilde{v}}$	.3667	σ^2_{ξ}	.0168
	(.0323)		(.0058)
θ	-.1775	λ	-.2155
	(.0394)		(.0281)
Ratio	.3004		

Notes: In Panel A the covariances are computed for the matched sample, and using all years pooled. Panel B reports the estimation of the sensitivity to permanent shocks, α, and the sensitivity to transitory shocks, β. For both regressions the dependent variable is $\Delta\omega_{ijt}$ and the explanatory variable is $\Delta\epsilon_{jt}$. The estimation procedure and specific instruments used in each regression are explained in the text. The $F-test$ refers to the first-stage regression. Standard errors are reported in parentheses; p-values in brackets. In Panel C, the Ratio is defined in the text.

trolling for error correlation within firms.[6] In each regression the specific instruments are defined for $k = 1, .., 9$. For both regressions, a likelihood-ratio test rejects the null that the extra powers of the instruments are redundant.[7] The overidentifying restriction tests do not reject the validity of instruments used in both regressions, and from the F-test we conclude that the instruments used in each regression are jointly significant. Finally, we performed the exogeneity test for $\Delta\epsilon_{jt}$ based on the difference in the Hansen-Sargan statistic between a model where it is assumed exogenous and our alternative model where we take it as endogenous. The test rejects the null that $\Delta\epsilon_{jt}$ is exogenous. This result implies that we also reject the equality between the sensitivity to both types of shocks.

We conclude from Panel B that workers' wages are not sensitive to transitory shocks on firms' performance, but they respond to firms' permanent shocks. The elasticity of wages to permanent shocks to firms' performance is 0.09 (compared to 0.07 in Guiso et al., 2005, for Italy).

Following the evidence provided by Altonji and Segal (1996), we estimated the different variance components using equally weighted minimum distance. Panel C reports the results. We can define the two variances associated with the shocks to sales as $\sigma_u^2 = \sigma_{\tilde{u}}^2/(1-\rho)^2$ and $\sigma_v^2 = (1+\theta^2)\sigma_{\tilde{v}}^2 + (\rho/(1-\rho))^2\sigma_{\tilde{u}}^2$. These are the variances of the permanent shock and the transitory shock, respectively. We estimate that σ_u^2 is 0.477, and σ_v^2 is 0.485, which amounts to a considerable variability. The moving average coefficient is about -0.18. All three estimates are statistically significant. For workers the variance of permanent shocks, σ_μ^2, is approximately 0.01 (statistically non-different from zero), while the variance of transitory shocks, σ_ξ^2, is 0.0168. The moving average parameter estimate is -0.22, and significant. These results are consistent with our analysis from Panel B. Our results also show that the different variances are considerably higher for firms than for workers.

To compute the portion of wage variability that can be attributed to firm's shock, the ratio $\sqrt{E\left\{\left[(\Delta\omega_{ijt})^2\right]|j\right\}}/\sqrt{E\left[(\Delta\omega_{ijt})^2\right]}$ is defined. We conclude that approximately 30% of the total variability in wages can be explained by firm-specific risk. For the Italian labour market, Guiso et al. conclude that this ratio is about 15%. In comparison with Italy, Portugal also presents much higher variances of the shocks for both sales and wages.

[6] In the permanent shock regression we clearly reject the null hypothesis of homoskedastic error terms, which justifies the use of GMM. For example, the Pagan and Hall test discussed in Baum et al. (2003) has a $p-value$ of 0.0148. For the transitory shock the evidence on heteroskedasticity is mixed. However, since our sample is large enough for asymptotic results to be valid, and given that IV gives inconsistent inference results if errors are in fact heteroskedastic, we adopted a conservative strategy and implemented the GMM procedure also in this case. The following conclusions on transitory shocks are not changed if we use generalized IV instead of GMM.

[7] The $p-value$ of the tests is always below 0.001.

Combining the evidence gathered so far, we conclude that Portuguese firms provide less insurance to their workers, when compared to Italian firms, a result in line with the high wage flexibility pointed out by studies on Portugal.

5.8 Forces shaping wage flexibility at the firm level

We now turn to the analysis of heterogeneity in insurance provision by firms. We consider different factors identified in the theoretical literature as shaping wage flexibility at the firm level. First of all, firms may be subject to institutional constraints. As argued by Faggio and Nickell (2005), national pay bargaining may insulate wages from firm idiosyncratic shocks. A similar role can be played by the minimum wage legislation, since firms with a large share of their workforce on minimum wage will have part of their wage policy set by the Government based on nation-wide trends. Firms that operate in more than one industry or region may be more able to diversify risk. On the contrary, a higher risk of going bankrupt will reduce the firm possibility to provide wage insurance. We consider also the occupation of the worker, with a dummy variable for managers meant to proxy two factors: the sensitivity of firm output to worker effort, with the wages of crucial workers more closely linked to firm performance, and therefore subject to less insurance provision; the capacity of the worker to bear risk, with managers likely to have more wealth and more access to financial markets where to diversify risk, and larger expertise in financial issues. The possibility of monitoring output has been pointed out as another factor that reduces the degree of insurance provided by the firm. Indeed, if the firm could monitor exactly the effort of the worker, it would not need to engage in a wage contract. Higher precision of the signal on the agent's effort will lead to less insurance (Guiso et al., 2005) have computed the noise on performance as the variability over time in the performance of the firm).

Empirically, we implement the heterogeneity analysis by extending the estimations in Panel B of Table 5.5 such that interactions of the explanatory variable $\Delta\epsilon_{jt}$ with firm and worker attributes are added. The dependent variable remains $\Delta\omega_{ijt}$. The results are reported in table 5.6, where $Manager$ is a dummy variable equal to one if he worker is a manager and *Decent. barg.* equals one if the worker is covered by firm-level bargaining, as opposed to a massive collective bargaining agreement.[8] $Bankruptcy$ is the threat of bankruptcy[9], $NInd$ is the number of industries in which the firm operates, $FSize$ stands

[8]Worker covered by a firm-level agreement or collective bargaining agreement (which involves a restricted group of firms, not organized into an employer association), as opposed to *collective bargaining contracts*, which often cover a whole industry, or the mandatory regime imposed by the Government.

[9]The percentage of firms that goes bankrupt in a given year and detailed region.

Table 5.6: Insurance heterogeneity

	Permanent shock	Transitory shock
$\Delta\epsilon_{jt}$	-.1628	-.0506
	(.1016)	(.0076)
	[.0281]	[.3192]
$\Delta\epsilon_{jt} * Manager$	.0194	-.0039
	(.0054)	(.0029)
	[.0157]	[.4370]
$\Delta\epsilon_{jt} * SDsales$	-.0078	.0023
	(.0081)	(.0013)
	[.0990]	[.4230]
$\Delta\epsilon_{jt} * Bankruptcy$	.0215	.0021
	(.0110)	(.0004)
	[.0298]	[.3736]
$\Delta\epsilon_{jt} * Foreign$	.0149	.0124
	(.0239)	(.0023)
	[.0380]	[.4229]
$\Delta\epsilon_{jt} * FSize$	-.0050	.0018
	(.0081)	(.0012)
	[.0486]	[.3717]
$\Delta\epsilon_{jt} * NInd$	.0101	.0063
	(.0245)	(.0041)
	[.0289]	[.4130]
$\Delta\epsilon_{jt} * Shareminw$	-.0380	.0095
	(.0229)	(.0410)
	[.3266]	[.1754]
$\Delta\epsilon_{jt} * Decent.barg.$	-.0017	.0171
	(.0298)	(.0030)
	[.0357]	[.5210]
Observations	25604	54873
J-test: $p - value$	.4309	.5447

Notes: The dependent variable is $\Delta\omega_{ijt}$. The instruments used in each regression are explained in the text. Robust standard errors reported in parentheses; Shea's (1997) partial R^2 in brackets. We account for within firm correlation of residuals. We report the J-test for the validity of the instruments.

for (log of) firm employment, and $Foreign$ is a dummy variable for the foreign origin of the capital; $SDsales$ represents the volatility of firm sales[10], and $Shareminw$ is the share of workers in the firm earning the national minimum wage.

To estimate these regressions we implemented once again the GMM procedure used in Panel B of Table 5.5, and define the extra instruments as the previous instruments interacted with the new variables. Note that different sets of instruments correspond to estimation results in columns 1 and 2 in that table, allowing for the estimation of heterogeneity in sensitivity to permanent and temporary shocks, respectively. The validity of each set of instruments is not rejected in both regressions, as indicated by the J-test. Since we have multiple endogenous regressors, Shea's (1997) partial R^2 are reported for each regressor used in the two regressions.

Results indicate that firms with a larger share of their workforce earning the minimum wage are less able to translate permanent shocks in product demand into wage changes. Indeed, the minimum wage is set at the national level by Government regulation, taking into explicit account aggregate trends such as the overall economy inflation rate. However, when faced with transitory shocks, firms with different shares of minimum wage workers do not react differently in terms of wage insurance. The level at which collective bargaining takes place also has an impact on the degree of insurance provided by the firm when faced with transitory shocks. More decentralized bargaining regimes are associated with less insurance, as opposed to massive collective wage setting agreements, which constraint the capacity of the firm to reflect demand shocks on wage changes.

Managers are less insured against permanent shocks than the rest of the workforce. This could be due to the fact that they may receive performance pay that links wages directly to the results of the company. Moreover, managers can be expected to be less risk-averse than other workers and as such would not have to be given the same level of insurance to exert effort. However, managers and workers with other occupations receive equal protection against transitory shocks.

Firms with a higher threat of bankruptcy are, as expected, less able to provide wage insurance and more constrained to reflect changes in product markets into changes in wages. That holds both for transitory and permanent shocks. Firms with higher variability in their sales offer less insurance against transitory changes in their performance. When faced with higher uncertainty in product markets, firms are bound to reflect more of the change in sales on wages, in the short run. Foreign firms provide less insurance to transitory shocks.

[10] Measured by the standard deviation of logarithm of sales for the years under analysis.

5.9 Conclusion

The impact of product market uncertainty on workers wages has been evaluated, relying on data of remarkable quality to estimate dynamic panel data models. Results point to the rejection of the full insurance hypothesis. Workers' wages respond to permanent shocks to firm performance, whereas they are not sensitive to transitory shocks. Based on the comparable analysis made for Italy by Guiso et al., we conclude that Portuguese firms provide less insurance to their workers. The higher responsiveness of wages to shocks at the firm level corroborates evidence previously reported on the high degree of wage flexibility in Portugal, when evaluated as the responsiveness of wages to macroeconomic conditions.

Another aim of the analysis was to check the impact of labour market regulations on the extent to which firms translate idiosyncratic shocks in product markets into shocks to the wages paid. We found that the national minimum wage and collective bargaining are indeed associated with the extent of wage insurance provided by the firm. Firms with a larger share of their workforce earning the minimum wage are less able to translate permanent shocks in product demand into wage changes. Also, massive collective wage setting agreements constraint the capacity of the firm to reflect idiosyncratic demand shocks into wage changes. This would be consistent with a corporatist wage setting view of the labour market, according to which the major role of these institutions would be to promote a smooth adjustment of wages to another type of shocks, those at the aggregate level.

Appendices to Chapter 5

5.A Checks on the consistency of data

After merging the worker data across years, inconsistencies were identified if the worker gender or date of birth was reported changing, or if the highest schooling level achieved was reported decreasing over time. In that case, the information reported over half the times has been taken as the correct one[11] (0.8%, 2.3%, 5.2% of the observations have been corrected, respectively for gender, birth date and education). Workers with inconsistent data after the introduction of the previous corrections were dropped. The whole information on the worker was dropped, whichever the incorrect number of observations identified (1.7%, 1.1%, and 4.3% of the observations, respectively for gender, birth date and schooling). Workers with missing age or schooling after the introduction of the previous corrections were dropped (respectively 0.7% and 1.7% of the observations, corresponding to 2.1% and 2% of the workers).

5.B Constraints imposed

The analysis focuses on workers and firms in manufacturing and services private sector in mainland Portugal.

On the worker side, we have retained wage-earners working full-time, aged 18 to 65, whose wage is not below the national minimum wage[12] (which led to dropping 20%, 2%, and 3% of the data set, respectively). Outliers in wage growth have been dropped,[13] which corresponded to a very small share of the data base, 0.03%. Workers observed just once in the database cannot be considered in the estimation of the models used (and thus 5% have been dropped). This is the full set of workers, which comprises over ten million observations. Due to the large size of the full data set it was not feasible to run the worker computations on the full data set and we have therefore used a 2 percent random sample of workers (keeping all the yearly observations for the selected workers). Descriptive statistics on this sample, comprising 205,352 yearly observations on 42,008 workers, are presented in table 5.7.[14]

[11] Note that this requirement is more demanding than just considering the modal value as the accurate one.

[12] May drop apprentices and handicapped workers.

[13] Log difference in real wages either greater than 2 or smaller than -.5.

[14] The dynamics in the models under estimation determine that a smaller number of individuals will be considered in the regressions.

On the firm side, we have kept firms operating full-year, and whose sales are not missing or outlier[15] (thus dropping 3%, 9%, and 0.2% of the firms, respectively).[16] Firms that were ever larger than 20 workers have been kept for analysis, since they are more likely to be run in entrepreneurial terms. Given the very small size structure of the firms in the Portuguese economy, this led to keeping 12% of the firms. The set of firms under analysis comprises 131,118 yearly observations on 18,368 firms. Descriptive statistics are reported in table 5.8.[17]

5.C Descriptive statistics

Gross monthly earnings were computed as $monthw = bw + sen + reg$,where bw stands for base-wage, sen are seniority-indexed components of pay, and reg are other regularly paid components. Wages were deflated using the Consumer Price Index.

Table 5.7: Descriptive statistics on workers

Variable	**Mean**	**Std. Dev.**
Log real monthly wage (PTE)	11.63	0.50
Age	36.2	10.91
Gender (female)	0.39	
Education		
4 years	0.46	
6 years	0.22	
9 years	0.13	
High School	0.14	
University	0.05	
Occupation		
managers	0.02	
professionals	0.02	
middle managers, technicians	0.09	
administrative	0.15	
service, sales	0.11	
skilled	0.27	

Continued on next page...

[15] Log difference in real sales either greater than 5 or smaller than -5.

[16] Firms in the first few months of their existence, not yet one year, were excluded, to avoid capturing sales fluctuations that are due to part-year operation.

[17] The dynamics in the models under estimation determine that a smaller number of firms will be considered in the regressions.

... table 5.7 continued

Variable	Mean	Std. Dev.
machine operators, assembly	0.14	
unskilled	0.15	
unknown	0.05	
Industry		
food, beverage, tobacco	0.05	
textiles	0.17	
wood	0.04	
chemicals	0.05	
mineral products	0.15	
construction	0.10	
trade	0.21	
restaurants, hotels	0.05	
transport, communications	0.04	
banking, insurance, business services	0.09	
other services	0.05	
Region		
North Coast	0.34	
Center Coast	0.16	
Lisbon	0.4	
Inland	0.08	
Algarve	0.03	
Type of collective bargaining agreement		
Decentralized	0.06	
Massive	0.94	
N	205352	

Table 5.8: Descriptive statistics on firms

Variable	**Mean**	**Std. Dev.**
Log real sales (1000 PTE)	12.93	1.45
Number workers in firm	58.22	170.8
Number of industries in firm	1.09	0.38
Share firms bankrupt in province	0.09	0.04
Variability firm sales over time: sd log real sales	0.5	0.51
Share of workers earning the minimum wage	0.03	0.11
Industry		
food, beverage, tobacco	0.05	
textiles	0.19	
wood	0.05	
chemicals	0.06	
mineral products	0.15	
construction	0.11	
trade	0.2	
restaurants, hotels	0.04	
transport, communications	0.04	
banking, insurance, business services	0.06	
other services	0.05	
Region		
North Coast	0.34	
Center Coast	0.18	
Lisbon	0.37	
Inland	0.08	
Algarve	0.03	
Origin of capital		
national	0.94	
foreign	0.06	
N	131118	

Chapter 6

Conclusion

This thesis is the compilation of four empirical studies on education, growth and labour economics. Education measurement error issues, the education of neighbouring countries, and the implication of both on growth regression estimation results are at the centre of the analyses in Chapters 2 and 3. In Chapters 4 and 5 we focus on labour economics topics. Specifically, we study gender discrimination and wage insurance provided by the firm, taking as an example the Portuguese labour market. Two main reasons justify the choice of the Portuguese case: (i) the richness of the data set available with information on matched employer-employee, which puts it among unique data sets used to analyse the labour market, and (ii), the Portuguese integration in the European Union in the last 20 years, which makes it a good case study to discuss and confront the results available in the literature. The main results of each chapter can be summarized as follows.

The starting point for Chapter 2 was the observation that there is a systematic difference in the education level between census data and observations constructed from enrolment data, in the education data provided by Barro and Lee (2001). A methodology was suggested for taking into account such measurement error, and its results point to an underestimation of education of about 0.2 years of education for every 5 years. In a regression context, although standard attenuation bias arguments suggest that using a corrected version of the data would lead to a higher coefficient, our growth regressions reveal the opposite. Our argument is that understatement of the average change in education helps explaining the higher coefficient obtained with the data released by Barro and Lee, when compared with results obtained for the corrected education data we build.

In face of the evidence, we have proposed an alternative explanation for the known difference between regressions based on 5 and on 10 year first–differences. A higher incidence of census observations in the 10 year data, when compared with the 5 year span data, is associated with a lower spurious variation in education, which helps explain why the coefficient of education in the growth regression is higher when we use data with larger

time spans.

The computations based on our corrected data show an immediate return to education bounded between 4.2% and 6.5%, while its long-run counterpart varies between 54% and 59%. These long-run returns to education should be compared with the time it takes to fully observe them. The half-life of this return is bounded between 75 and 99 years, which makes the span of the data currently available too short to identify with precision long-run returns to education. For example, using the data set released by Cohen and Soto (2001) we have confirmed the positive contemporaneous returns to education, but a smaller long-run return of 49% was uncovered. The overall evidence supports both human capital and externalities explanations for the role of education in economic growth. Additional estimations using dynamic panel data methods indicate that our main estimation results are statistically consistent.

A limitation of the analysis is the simplicity of the growth regressions used to test the role of education on economic growth. This might raise some concern on the reliability of the estimated returns to education. Although we believe that the models estimated capture the essence on the question under analysis, an extension to consider in future research is to analyse how the corrected education measure would perform in more complete growth regressions.

In Chapter 3 we have analysed education spillovers across countries. By using bilateral trade to quantify the interaction between countries, we have built a measure of neighboring education for each country. Consistent with the findings in the previous chapter, countries' investment in education have both contemporaneous and long-run effects in their growth process. This positive effect also spills-over to their neighbours. Returns to education exceed the private returns identified in the literature, and significant externalities of education are revealed in the long-run. The internal short-run return to education is about 6% to 7%, while the internal return in the 40-year interval of our analysis is above 11%. Within this time frame, the returns to external education can be as high as 17%. By controlling for neighboring education we obtain a higher return to education when compared with the previous chapter. In the long-run, the return to own education is above 34%, while the total return is at least 66%. Depending on the weights we use, external returns can achieve 52% in the long-run. Once again, we have to account for the fact that the half-life period lies above 65 years.

We have illustrated those results by means of a series of examples. In particular, we have computed the returns to education in three countries, namely: the US, The Netherlands, and the UK. The US, with an average education of 12 years in 2000, has lower returns to education than, for example, The Netherlands and the UK, who have

lower average levels of education. By accounting for the time frame and the relative position of each country in the world distribution of education, we have concluded that cross-country evidence is compatible with the results of the literature on private returns to education.

Results on neighbouring education are less clear when we use non-standardized weights. Although having "good" neighbours, as measured by their specific effect, has a positive impact on economic performance, the final result presented in Chapter 3 reveals that the role of neighbouring education is not easily identified in this context. Further research is needed in order to disentangle "good" neighbours from neighbours with a "good" education.

From a policy point of view, we conclude that investments in education have an important return in terms of economic performance, both in the short and the long run. In spite of the measurement error problems commonly associated with growth regressions and education data, our analysis indicates that investments in education have both level and growth effects in the countries' economic performance. Furthermore, accounting for interactions between countries reinforces the conclusion that overall returns to education are considerable, and seem to justify the investment in education that has been observed across countries.

In order to complement the analysis by linking regional and private returns to education within countries or economic areas, we could apply this type of analysis, for instance, to European regions. This way, we would reduce data quality problems, and the analysis would benefit from a higher degree of homogeneity of the units of observation. An important extension of the analysis is to compare the returns to education to the ones associated to alternative investments; e.g. on health, infrastructures, good institutions, or innovation. Such computations are relevant to evaluate whether further investments in education are the efficient way to foster economic growth, once a certain level of investment is achieved. Given the simplicity of the gravity model used as the basis of the definition of the weights, it is important to consider in future analysis a more complex formulation of the gravity equation. The strong results we observe on returns to education suggest that we also should replicate our analysis using more extensive growth regressions, in order to check for potential problems related to misspecification.

In chapters 4 and 5 the analysis focuses on the Portuguese labour market. The results in Chapter 4 show a high degree of systematic gender segregation across establishments, which remained relatively stable between 1985 and 1999. The total gender wage gap is high and persistent, being bounded between 25% and 29%. We conclude that a higher proportion of females in the establishment lowers females' wages, having the opposite effect

on males' wages. The establishment femaleness accounts for 10% to 25% of the wage gap. In order to capture heterogeneity across firms, we have extended the original analysis by including a set of firm specific effects. Although the contribution of the share of females to the total wage gap becomes smaller, it is still substantial, and reveals a rising trend. The overall results show that both the taste-based model of employer behavior and the theory of sorting of workers across establishments based on their productivity are not rejected by the data, although an eye catching result is that gender discrimination mechanisms became less important, as the apparent compensation for male workers has decreased over time. Female disadvantages, compared to male advantages, explain a larger portion of the wage gap. This is particularly relevant when we look at the proportion of females in the establishment.

Governmental policies aiming at enforcing equal pay and lower gender segregation levels across firms and establishments would improve women outcomes in the labour market. Specific policies should target women's education and training, particularly of less skilled women. Such measures could counter discrimination and lead to the decline of the wage gap.

Finally, in Chapter 5, we analysed wage flexibility in the form of wage insurance provided by firms. Estimation results point to the rejection of the full insurance hypothesis. Permanent shocks to firm performance are reflected in workers' wages, which are not sensitive to transitory shocks. We found an elasticity of wages to permanent shocks in firms' performance of about 0.09. Interesting is the fact that managers receive similar protection against transitory shocks as co-workers in other occupations, but they are less insured against permanent shocks. Higher sales' variability and being controlled by foreign firms are associated with lower insurance against transitory shocks. A higher threat of bankruptcy reduces the possibility of the firm offering wage insurance to its workers. Contrasting our results with the ones available in the literature, and picking the comparable analysis made for Italy, we conclude that Portuguese firms provide less insurance. This is consistent with the high wage flexibility associated with the Portuguese labour market.

When we look at the impact of labour market regulations on wage flexibility, two main patterns emerge. First, we find that wages are less sensitive to permanent demand shocks in firms with larger proportions of minimum wage workers. Second, the existence of a centralized wage agreement also constrain the capacity of the firm to reflect idiosyncratic demand shocks into wage changes.

A shortcoming in our analysis is the fact that we do not consider employment effects of shocks to firms' performance. The joint analysis of wages and firing and hiring decisions,

accounting for firms' idiosyncratic shocks, is the natural extension to the current analysis.

Summing up, throughout this thesis we have discussed aspects within the economics of education, economic growth and labour economics. On the one hand, this thesis is an attempt to throw additional light in the puzzling relation between education and economic performance. In particular, we look at the role of neighbouring education in the economic growth process by using a new approach to further understand how education may impact on economic outcomes. On the other hand, we have replicated and extended previous studies on gender discrimination and insurance provision by the firm at the example of the Portuguese labour market. More research is needed, however, to address all the unanswered questions raised in the current thesis.

References

Abowd, John A. and Thomas Lemieux (1993), 'The effects of product market competition on collective bargaining agreements: The case of foreign competition in Canada', *Quarterly Journal of Economics* **108**(4), 983–1014.

Acemoglu, Daron and Joshua Angrist (2000), How large are human capital externalities? Evidence from compulsory schooling laws, *in* 'NBER Macroeconomics Annual', Vol. 15, MIT Press, Cambridge, pp. 9–59.

Aghion, Philippe and Peter Howitt (1998), *Endogenous Growth Theory*, MIT Press, Cambridge, MA.

Altonji, Joseph G. and Lewis M. Segal (1996), 'Small-sample bias in GMM estimation of covariance structures', *Journal of Business and Economic Statistics* **14**(3), 353–367.

Arellano, Manuel (2003), *Panel Data Econometrics*, Oxford University Press, New York.

Arellano, Manuel and Olimpia Bover (1995), 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics* **68**(1), 29–51.

Arellano, Manuel and Stephen Bond (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies* **58**(2), 277–297.

Azariadis, Costas (1975), 'Implicit contracts and underemployment equilibria', *Journal of Political Economy* **83**(6), 1183–1202.

Baily, Martin Neil (1974), 'Wages and employment under uncertain demand', *Review of Economic Studies* **41**(1), 37–50.

Barro, Robert J. and Jong-Wha Lee (1993), 'International comparisons of educational attainment', *Journal of Monetary Economics* **32**(3), 363–394.

Barro, Robert J. and Jong-Wha Lee (1996), 'International measures of schooling years and schooling quality', *American Economic Review* **86**(2), 218–223.

Barro, Robert J. and Jong-Wha Lee (2001), 'International data on educational attainment: Updates and implications', *Oxford Economic Papers* **53**(3), 541–563.

Barro, Robert J. and Xavier Sala-i-Martin (1999), *Economic Growth*, MIT Press, Cambridge, Massachusetts.

Baum, Christopher F., Mark E. Schaffer and Steven Stillman (2003), 'Instrumental variables and GMM: Estimation and testing', *The Stata Journal* **3**(1), 1–31.

Bayard, Kimberly, Judith Hellerstein, David Neumark and Kenneth Troske (2003), 'New evidence on sex segregation and sex differences in wages from matched employer-employee data', *Journal of Labor Economics* **21**(4), 887–922.

Beaudry, Paul and John DiNardo (1991), 'The effects of implicit contracts on the movement of wages over the business cycle: Evidence from micro data', *Journal of Political Economy* **99**(4), 665–688.

Becker, Gary S. (1971), *The Economics of Discrimination*, University of Chicago Press, Chicago.

Benhabib, J. and M. Spiegel (1994), 'The role of human capital in economic development: Evidence from aggregate cross-country data', *Journal of Monetary Economics* **34**(2), 143–173.

Blanchflower, David G., Andrew J. Oswald and Peter Sanfey (1996), 'Wages, profits and rent-sharing', *Quarterly Journal of Economics* **111**(1), 227–251.

Blundell, Richard and Stephen Bond (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics* **87**(1), 115–143.

Boisso, Dale, Kathy Hayes, Joseph Hirschberg and Jacques Silber (1994), 'Occupational segregation in the multidimensional case', *Journal of Econometrics* **61**(1), 161–171.

Bond, Stephen R., Anke Hoeffler and Jonathan R.W. Temple (2001), 'GMM estimation of empirical growth models', Centre for Economic Policy Research Discussion Paper No. 3048.

Bowsher, Clive G. (2002), 'On testing overidentifying restrictions in dynamic panel data models', *Economics Letters* **77**(2), 211–220.

Branco, Fernando and António Mello (1992), Why are wages in Portugal lower than elsewhere in EEC?, *in* J. Amaral and D. Lucena, eds, 'The Portuguese Economy Towards 1992', The Portuguese Economy Towards 1992, Kluwer, Boston, pp. 131–151.

Bruno, Giovanni (2005*a*), 'Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models', *Economics Letters* **87**(3), 361–366.

Bruno, Giovanni (2005*b*), 'Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals', *Stata Journal* **5**(4), 473–500.

Buckley, John E. (1971), 'Pay differences between men and women in the same job', *Monthly Labor Review* **64**(11), 36–39.

Bun, Maurice and Jan Kiviet (2003), 'On the diminishing returns of higher-order terms in asymptotic expansion of bias', *Economics Letters* **79**(2), 145–152.

Cappellari, Lorenzo (1999), 'Minimum distance estimation of covariance structures'. Paper presented at the Conference of the Royal Statistical Society, London, UK, May 17.

Cardoso, Ana Rute (1998), 'Earnings inequality in Portugal: High and rising?', *Review of Income and Wealth* **44**(3), 325–343.

Cardoso, Ana Rute (1999), 'Firm's wage policies and the rise in labor market inequality: The case of Portugal', *Industrial and Labor Relations Review* **53**(1), 87–102.

Cardoso, Ana Rute and Miguel Portela (2005), 'The provision of wage insurance by the firm: Evidence from a longitudinal matched employer-employee dataset'. IZA Discussion Paper 1865, Bonn.

Cardoso, Ana Rute and Pedro Portugal (2005), 'Contractual wages and the wage cushion under different bargaining settings', *Journal of Labor Economics* **23**(4), 875–902.

Carrington, William J. and Kenneth R. Troske (1995), 'Gender segregation in small firms', *Journal of Human Resources* **30**(3), 503–533.

Carrington, William J. and Kenneth R. Troske (1997), 'On measuring segregation in samples with small units', *Journal of Business and Economic Statistics* **15**(4), 402–409.

Carrington, William J. and Kenneth R. Troske (1998), 'Sex segregation in the u.s. manufacturing', *Industrial and Labor Relations Review* **51**(3), 445–464.

Christofides, Louis N. and Andrew J. Oswald (1992), 'Real wage determination and rent-sharing in collective bargaining agreements', *Quarterly Journal of Economics* **107**(3), 985–1002.

Cohen, Daniel and Marcelo Soto (2001), 'Growth and human capital: Good data, good results', OECD Development Centre Technical Paper No. 179, Paris.

Conley, Timothy G. and Ethan Ligon (2002), 'Economic distance and cross-country spillovers', *Journal of Economic Growth* **7**(2), 157–187.

Cotton, Jeremiah (1988), 'On the decomposition of wage differentials', *Review of Economics and Statistics* **70**(2), 236–243.

Davidson, Russell and James G. MacKinnon (2004), *Econometric Theory and Methods*, Oxford University Press, New York.

De la Fuente, Angel and Rafael Doménech (2002), 'Human capital in growth regressions: How much difference does data quality make? An update and further results', CEPR Discussion Paper no. 3587, London.

De la Fuente, Angel and Rafael Doménech (2006), 'Human capital in growth regressions: How much difference does data quality make?', *Journal of the European Economic Association* **4**(1), 1–36.

Devereux, Paul J. (2005), 'Do employers provide insurance against low frequency shocks? industry employment and industry wages', *Journal of Labor Economics* **23**(2), 313–340.

Doornik, J.A., M. Arellano and S. Bond (2002), *Panel Data Estimation Using DPD for Ox*, Oxford, London. www.nuff.ox.ac.uk/Users/Doornik.

Estevão, Marcello and Stacey Tevlin (2003), 'Do firms share their success with workers? the response of wages to product market conditions', *Economica* **70**(280), 597–617.

Faggio, Guilia and Stephen Nickell (2005), 'The responsiveness of wages to labour market conditions in the UK', *Labour Economics* **12**(5), 685–696.

Flückiger, Yves and Jacques Silber (1999), *The Measurement of Segregation in the Labor Force*, Physica-Verlag, Heidelberg.

Gamber, Edward (1988), 'Long-term risk-sharing wage contracts in an economy subject to permanent and temporary shocks', *Journal of Labor Economics* **6**(1), 83–99.

Glaeser, Edward L., Bruce I. Sacerdote and José A. Scheinkman (2003), 'The social multiplier', *Journal of the European Economic Association* **1**(2-3), 345–353.

Goldin, Claudia (1990), *Understanding the Gender Gap*, Oxford University Press, Oxford.

Gordon, Donald F. (1974), 'A neo-classical theory of keynesian unemployment', *Economic Inquiry* **12**(4), 431–459.

Gradstein, Mark and Moshe Justman (2002), 'Education, social cohesion, and economic growth', *American Economic Review* **92**(4), 1192–1204.

Groshen, Erica (1991), 'The structure of the female/male wage differential: Is it who you are, what you do, or where you work?', *Journal of Human Resources* **26**(3), 457–472.

Guiso, Luigi, Luigi Pistaferri and Fabiano Schivardi (2005), 'Insurance within the firm', *Journal of Political Economy* **113**(5), 1054–1087.

Gyimah-Brempong, Kwabena, Rudy Fichtenbaum and Gregory Willis (1992), 'The effects of college education on the male-female wage differential', *Southern Economic Journal* **58**(3), 790–804.

Harmon, Colm, Hessel Oosterbeek and Ian Walker (2003), 'The returns to education: Microeconomics', *Journal of Economic Surveys* **17**(2), 115–155.

Harris, Milton and Bengt Holmstrom (1982), 'A theory of wage dynamics', *Review of Economic Studies* **49**(3), 315–333.

Hartog, Joop, Pedro Pereira and José Vieira (2000), 'Inter-industry wage dispersion in Portugal', *Empirica* **27**(4), 353–364.

Hartog, Joop, Pedro Pereira and José Vieira (2001), 'Changing returns to education in Portugal during the 1980s and the early 1990s: OLS and quantile regression analysis', *Applied Economics* **33**(8), 1021–1037.

Heston, Alan, Robert Summers and Bettina Aten (2002), 'Penn World Table version 6.1'. Center for International Comparisons at the University of Pennsylvania (CICUP), October, Philadelphia.

Hildreth, Andrew K. G. and Andrew J. Oswald (1997), 'Rent-sharing and wages: Evidence from company and establishment panels', *Journal of Labor Economics* **15**(2), 318–337.

Holmstrom, Bengt (1981), 'Contractual models of the labor market', *American Ecoomic Review* **71**(2), 308–313.

Holzer, Harry J. and Edward B. Montgomery (1993), 'Asymmetries and rigidities in wage adjustments by firms', *Review of Economics and Statistics* **75**(3), 397–408.

Hutchens, Robert (1991), 'Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations', *Mathematical Social Sciences* **21**(1), 31–51.

Hutchens, Robert (2001), 'Numerical measures of segregation: Desirable properties and their implications', *Mathematical Social Sciences* **42**(1), 13–29.

Kiker, B.F. and Maria C. Santos (1991), 'Human capital and earnings in Portugal', *Economics of Education Review* **10**(3), 187–203.

Kremer, Michael (1993), 'The o-ring theory of economic development', *Quarterly Journal of Economics* **108**(3), 551–575.

Kremer, Michael and Eric Maskin (1996), 'Wage inequality and segregation by skill'. NBER working paper 5718, Cambridge, MA.

Krueger, Alan and Mikael Lindahl (2001), 'Education for growth: Why and for whom?', *Journal of Economic Literature* **39**(4), 1101–1136.

Kyriacou, George (1991), 'Level and growth effects of human capital: A cross-country study of the convergence hypothesis'. New York University Economic Research Report: 91-26.

Linders, Gert-Jan M. and Henri L.F. de Groot (2006), 'Estimation of the gravity equation in the presence of zero flows', Tinbergen Institute Discussion Paper 2006-072/3, Amsterdam-Rotterdam.

Lucas, Robert E., Jr. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* **22**(1), 3–42.

Mankiw, N. Gregory, David Romer and David N. Weil (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics* **107**(2), 407–438.

Manski, Charles (1993), 'Identification of endogenous social effects: The reflection problem', *Review of Economic Studies* **60**(3), 531–542.

McNulty, Donald J. (1967), 'Differences in pay between men and women workers', *Monthly Labor Review* **90**(12), 40–43.

Moreno, Ramon and Bharat Trehan (1997), 'Location and the growth of nations', *Journal of Economic Growth* **2**(4), 399–418.

Moretti, Enrico (2004*a*), 'Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data', *Journal of Econometrics* **121**(1-2), 175–212.

Moretti, Enrico (2004*b*), 'Workers' education, spillovers, and productivity: Evidence from plant-level production functions', *American Economic Review* **94**(3), 656–690.

Mulligan, Casey B. and Xavier Sala-i-Martin (2000), 'Measuring aggregate human capital', *Journal of Economic Growth* **5**(3), 215–252.

Nehru, V., E. Swanson and A. Dubey (1995), 'A new database on human capital stocks in developing and industrial countries: Sources, methodology and results', *Journal of Development Economics* **46**(2), 379–401.

Nelson, Richard and Edmund Phelps (1966), 'Investment in humans, technology diffusion and economic growth', *American Economic Review* **56**(2), 69–75.

Neumark, David (1988), 'Employers' discriminatory behavior and the estimation of wage discrimination', *Journal of Human Resources* **23**(3), 279–295.

Nickell, Stephen and Sushil Wadhwani (1990), 'Insider forces and wage determination', *Economic Journal* **100**(401), 496–509.

Oaxaca, Ronald (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review* **14**(3), 693–709.

OECD (Organisation for Economic Co-Operation and Development) (1992), *OECD Economic Surveys, Portugal, 1991/92*, OECD, Paris.

OECD (Organisation for Economic Co-Operation and Development) (1994*a*), *Economic Surveys, Portugal, 1993-1994*, OECD, Paris.

OECD (Organisation for Economic Co-Operation and Development) (1994*b*), *OECD Employment Outloook 1994*, OECD, Paris.

OECD (Organisation for Economic Co-Operation and Development) (1999), *OECD Employment Outlook 1999*, OECD, Paris.

Pfeffer, Jeffrey and Alison Davis-Blake (1987), 'The effect of the proportion of women on salaries: The case of college administrators', *Industrial and Labor Relations Review* **31**(1), 1–24.

Portela, Miguel, Rob Alessie and Coen Teulings (2004), 'Measurement error in education and growth regressions', Tinbergen Institute Discussion Paper 2004-040/3, Amsterdam-Rotterdam.

Portugal. Ministério da Segurança Social e do Trabalho (1985-2000), 'Quadros de pessoal'. Data in magnetic media.

Pritchettt, Lant (2001), 'Where has all the education gone?', *The World Bank Economic Review* **15**(3), 367–391.

Reilly, Kevin T. and Tony S. Wirjanto (1999), 'Does more mean less? The Male/Female wage gap and the proportion of females at the establishment level', *Canadian Journal of Economics* **32**(4), 906–929.

Rose, Andrew K. (2004), 'Do we really know that the WTO increases trade?', *American Economic Review* **94**(1), 98–114.

Rosen, Sherwin (1985), 'Implicit contracts: A survey', *Journal of Economic Literature* **23**(3), 1144–1175.

Shea, John (1997), 'Instrument relevance in multivariate linear models: A simple measure', *Review of Economics and Statistics* **79**(2), 348–352.

Sianesi, Barbara and John Van Reenen (2003), 'The returns to education: Macroeconomics', *Journal of Economic Surveys* **17**(2), 157–200.

Summers, Robert and Alan Heston (1991), 'The Penn World Table (mark 5): An expanded set of international comparisons, 1950-1988', *Quarterly Journal of Economics* **106**(2), 280–315.

Temple, Jonathan R.W. (1999), 'A positive effect of human capital on growth', *Economics Letters* **65**(1), 131–134.

Temple, Jonathan R.W. (2001), 'Generalizations that aren't? Evidence on education and growth', *European Economic Review* **45**(4-6), 905–918.

Teulings, Coen (1997), 'A new theory of corporatism and wage setting', *European Economic Review* **41**(3-5), 659–669.

Teulings, Coen and Thijs Van Rens (2006), 'Education, growth, and income inequality', *The Review of Economics and Statistics, forthcoming* .

Thomas, Jonathan and Tim Worrall (1988), 'Self-enforcing wage contracts', *Review of Economic Studies* **55**(4), 541–553.

Topel, Robert (1999), *Labor Markets and Economic Growth*, Orley Ashenfelter and David Card, eds, Amsterdam: North Holland, chapter 44, pp. 2943–2984.

Vieira, José A. Cabral (1999), The Evolution of Wage Structures in Portugal 1982-1992, PhD thesis, Tinbergen Institute, Amsterdam.

Vieira, José A. Cabral, Ana Rute Cardoso and Miguel Portela (2005), 'Gender segregagtion and the wage gap in Portugal: An analysis at the establishment level', *Journal of Economic Inequality* **3**(2), 145–168.

Weinberg, Bruce (2001), 'Long-term wage fluctuations with industry-specific human capital', *Journal of Labor Economics* **19**(1), 231–264.

Weiss, Yoram (1984), 'Wage contracts when output grows stochastically: The roles of mobility costs and capital market imperfections', *Journal of Labor Economics* **2**(2), 155–173.

Windmeijer, Frank (2005), 'A finite sample correction for the variance of linear efficient two-step GMM estimators', *Journal of Econometrics* **126**(1), 25–51.

Winter-Ebmer, Rudolf and Josef Zweimüller (1996), 'Immigration and the earnings of young native workers', *Oxford Economic Papers* **48**(3), 473–491.

Woessmann, Ludger (2003), 'Specifying human capital', *Journal of Economic Surveys* **17**(3), 239–270.

Yoon, Soohyun, Kenneth R. Troske and Peter Mueser (2003), 'Changes in gender segregation and women's wages in Korea'. Mimeo. University of Missouri-Columbia.

Samenvatting (Summary in Dutch)

De basis van dit proefschrift bestaat uit vier hoofdstukken die elk overeenkomen met een artikel. Elk hoofdstuk is een op zichzelf staande studie, onafhankelijk in motivatie, methodes en resultaten. De benadering door de verschillende hoofdstukken heen is voornamelijk een empirische en omvat zowel macro- als micro-economische perspectieven.

Het eerste deel van het proefschrift, dat de hoofdstukken 2 en 3 beslaat, concentreert zich op de rol die onderwijs speelt bij economische groei. De analyse wordt uitgevoerd met behulp van paneldata voor landen. De onderliggende vraag is of onderwijs een positief effect heeft op de economische groei van een land. Hoofdstuk 2 en 3 behandelen respectievelijk meetfouten en neveneffecten van onderwijs. In het tweede deel van het proefschrift wordt de aandacht verlegd naar een micro-economische analyse van de arbeidsmarkt, waarbij gebruik wordt gemaakt van een longitudinale verzameling gekoppelde werkgever-werknemergegevens voor Portugal. Hoofdstuk 4 analyseert de segregatie naar sekse en inkomensverschillen en hoofdstuk 5 heeft inkomensflexibiliteit en de verstrekking van loonsverzekering door bedrijven tot onderwerp. Hieronder volgt een korte samenvatting van de afzonderlijke vier hoofdstukken.

Hoofdstuk 2 is gebaseerd op Portela et al. (2004) en bespreekt het probleem van meetfouten in data over het onderwijsniveau in landen en het effect daarvan op groeiregressies. De *perpetual inventory* methode die gebruikt wordt door Barro en Lee (2001) om het onderwijsniveau van landen te meten, leidt tot een systematische meetfout. In dit hoofdstuk analyseren we het effect van zulke meetfouten op BNP-regressies. We vergelijken onze schattingen met die van Topel (1999) en Krueger en Lindahl (2001) en herhalen de analyse met de gegevens verstrekt door Cohen en Soto (2001). We tonen aan dat er een systematisch verschil in gemeten onderwijsniveau bestaat tussen censusgegevens en observaties afgeleid uit schoolinschrijvingsgegevens en bespreken een methode voor de correctie van deze meetfouten.

De resultaten gebaseerd op deze methode wijzen op een onderschatting van onderwijs van ongeveer 0,2 jaar op elke 5 jaar. De standaard *attenuation bias* zou suggereren dat het gebruik van gecorrigeerde data zal leiden tot een groter effect van onderwijs op

economische groei. De resultaten van onze schattingen laten echter het tegenovergestelde zien. Vervolgens bespreken we waarom de meetfout tot een overschatting leidt. In het licht van deze resultaten bieden we een alternatieve verklaring voor discrepanties tussen regressies gebaseerd op eerste verschillen voor intervallen van 5 en 10 jaar. Een hogere frequentie van censusobservaties in de 10-jaargegevens houdt verband met een kleinere schijnvariatie in onderwijs, in vergelijking met de 5-jaargegevens. Dit kan één verklaring zijn waarom de onderwijscoëfficiënt in groeiregressies hoger is wanneer de observaties een grotere tijdsspanne omvatten.

In de analyse beschouwen we korte- en lange-termijn effecten van onderwijs op economische groei, evenals de onderliggende halfwaardetijd, voor de paneldata met 5 en 10-jaars intervallen. We testen de robuustheid van de resultaten door een dynamische panelanalyse toe te passen op het vereenvoudigde groeimodel. De berekeningen gebaseerd op onze gecorrigeerde data laten een direct rendement van onderwijs tussen 4,2% en 6,5% zien, terwijl het lange-termijn effect varieert tussen 54% en 59%. Het lange-termijn effect van onderwijs zou vergeleken moeten worden met de tijd die nodig is om het volledig te kunnen observeren. De halfwaardetijd van dit effect ligt tussen 75 en 99 jaar, waardoor de reikwijdte van de huidige beschikbare gegevens tekort schiet om met precisie lange-termijn effecten van onderwijs te identificeren. Door gebruik te maken van een alternatieve dataset, samengesteld door Cohen en Soto (2001), hebben we de positieve directe effecten van onderwijs bevestigd, maar vonden we een lager lange-termijn effect van 49%. De globale bevindingen ondersteunen zowel het directe effect van menselijk kapitaal als externe effecten als verklaringen voor de bijdrage van onderwijs aan economische groei. Aanvullende schattingen gebaseerd op dynamische paneldata modellen wijzen erop dat onze belangrijkste resultaten consistent zijn.

In hoofdstuk 3 analyseren we externe effecten van onderwijs tussen landen. Door de aandacht te richten op landelijke gegevens proberen we globale neveneffecten van onderwijs te meten. We stellen een maat voor onderwijs van het buurland voor en gebruiken die om de rol van onderwijs in BNP-regressies te herwaarderen. We bespreken economische nabijheid van landen (een reflectie van de economische interactie tussen landen) en introduceren een definitie van economische nabijheid die wordt gebruikt in de rest van het hoofdstuk. We ontwikkelen een procedure om deze nabijheid te meten, gebaseerd op bilaterale handel. Vervolgens schatten we een dynamisch groeimodel dat zowel het onderwijs van het eigen land als dat van het buurland omvat. Daarnaast testen we de aanwezigheid van landspecifieke effecten, maar de resultaten van de hoofdspecificatie suggereren dat deze niet aanwezig zijn. We schatten het model door middel van niet-lineaire kleinste kwadraten. Dit stelt ons in staat een nabijheidsafnameparameter te schatten,

welke is gerelateerd aan de definitie van economische nabijheid.

In overeenstemming met de resultaten uit het voorgaande hoofdstuk hebben de investeringen van landen in onderwijs zowel onmiddellijke effecten als lange-termijn effecten op het groeiproces. Dit positieve effect werkt ook door naar de buurlanden. Het totale effect van onderwijs overtreft het particuliere effect dat in de literatuur wordt aangetoond en er worden significante neveneffecten van onderwijs gevonden op de lange termijn. Het interne korte-termijn effect van onderwijs ligt rond de 6 à 7%, terwijl het interne effect in het 40-jaarinterval van onze analyse hoger is dan 11%. Binnen dit tijdskader kan het effect van "extern onderwijs" 17% bereiken. Door het controleren van "aangrenzend onderwijs" vinden we een groter effect in vergelijking met het vorige hoofdstuk. Het lange-termijn effect van eigen onderwijs ligt boven de 34%, terwijl het totale effect ten minste 66% bedraagt. Afhankelijk van de gewichten die we gebruiken, kunnen externe effecten op de lange termijn 52% bereiken. Nogmaals, we moeten er rekening mee houden dat de halfwaardetijd boven de 65 jaar ligt. Resultaten voor aangrenzend onderwijs zijn minder duidelijk als we niet-gestandaardiseerde gewichten gebruiken.

Met betrekking tot beleid concluderen we dat investeringen in onderwijs een belangrijke bijdrage leveren aan economische prestaties, zowel op de korte termijn als op de lange termijn. Ondanks de problemen van meetfouten waar groeiregressie en onderwijsgegevens onder gebukt gaan, laat onze analyse zien dat onderwijs zowel niveau- als groeieffecten heeft op de economische prestaties van landen. Als we daarnaast controleren voor interacties tussen landen versterken de resultaten de conclusie dat de totale opbrengsten van onderwijs aanzienlijk zijn. Deze resultaten rechtvaardigen de investering in onderwijs die in de verschillende landen wordt waargenomen.

In hoofdstuk 4, gebaseerd op Vieira et al. (2005), analyseren we de trend in werknemerssegregatie op instellingsniveau en de impact hiervan op de lonen in Portugal over een periode van vijftien jaar. In dit hoofdstuk staan arbeidsmarktverschillen tussen mannen en vrouwen centraal: (i) wat is het niveau van segregatie naar sekse in organisaties op de Portugese arbeidsmarkt en hoe heeft dit zich in de tijd ontwikkeld?; (ii) wat is het effect van die segregatie op lonen?; en (iii), is dat effect voor mannen en vrouwen verschillend? Vanwege de gebruikelijke afwijkingen in de berekening van traditionele segregatiemetingen, construeren we een alternatieve maatstaf voor segmentatie naar sekse. We kwantificeren de mate van systematische segregatie als afstand tussen algemene segregatie en random segregatie. We gebruiken standaard loondecompositietechnieken om de impact van de samenstelling van de arbeidsbevolking op organisatieniveau op lonen vast te stellen.

De resultaten laten een hoge mate van systematische segregatie naar sekse op organ-

isatieniveau zien die relatief stabiel bleef tussen 1985 en 1999. De totale inkomenskloof tussen de seksen is hoog en hardnekkig en ligt tussen de 25 en 29%. We concluderen dat een hoger aandeel van vrouwen in de organisatie de lonen van vrouwen verlaagt, terwijl we het tegenovergestelde effect vinden voor de lonen van mannen. De ´vrouwelijkheid´ van de organisatie verklaart 10 tot 25% van de inkomenskloof. Om de verschillen tussen bedrijven te ondervangen, hebben we de oorspronkelijke analyse uitgebreid met een set van bedrijfsspecifieke effecten. Als we voor deze effecten controleren zien we weliswaar dat de bijdrage van het aandeel van vrouwen in de totale inkomenskloof kleiner wordt, maar dat deze nog steeds substantieel is en een stijgende lijn laat zien. Op basis van de resultaten kunnen zowel het voorkeursgebaseerde model van werkgeversgedrag als de theorie van werknemersselectie door organisaties op basis van hun productiviteit niet worden verworpen. Een opvallend resultaat is dat mechanismen van seksediscriminatie minder belangrijk zijn geworden. Vrouwelijke nadelen vergeleken met mannelijke voordelen verklaren een groter deel van de inkomenskloof. Dit is met name relevant als we kijken naar het aandeel van vrouwen in de organisatie.

Overheidsbeleid dat een gelijke beloning en een lager scheidingsniveau naar sekse in bedrijven en organisaties tot doel heeft, zou de positie van vrouwen op de arbeidsmarkt verbeteren. Specifiek beleid zou de opleiding en training van vrouwen tot doel moeten hebben, in het bijzonder van lager opgeleide vrouwen. Zulke maatregelen zouden discriminatie tegengaan en leiden tot een vermindering van de inkomenskloof.

Hoofdstuk 5 ten slotte, een herziene versie van Cardoso en Portela (2005), analyseert de verstrekking van inkomensverzekering door het bedrijf in de context van micro fundamenten voor inkomensflexibiliteit. We bestuderen de impact van productiemarktonzekerheid op lonen, waarbij we de volgende vragen stellen: (i) wat is de ontvankelijkheid van lonen voor schokken in bedrijfsresultaat?; (ii) welke bedrijfs- en werknemerskenmerken kunnen in verband worden gebracht met een hogere mate van inkomensflexibiliteit op microniveau? We bekijken met name de rol van arbeidsmarktbepalingen die de ontvankelijkheid van lonen voor bedrijfsfluctuaties beperken. We baseren ons eerst op Guiso et al. (2005), waarbij we dynamische modellen voor verkoop en lonen schatten om de gevoeligheid van lonen voor permanente en tijdelijke fluctuaties in het bedrijfsresultaat te bepalen. Daarna onderzoeken we de factoren die te maken hebben met hogere loonsflexibiliteit.

De resultaten wijzen erop dat de volledige verzekering hypothese verworpen moet worden. Permanente schokken in het bedrijfsresultaat worden weerspiegeld in werknemerslonen. Deze zijn echter niet gevoelig voor tijdelijke schokken. We hebben een loonelasticiteit ten opzichte van permanente fluctuaties in bedrijfsresultaat gevonden die ongeveer 0,09

bedraagt. Opvallend is het feit dat managers over een vergelijkbare verzekering tegen tijdelijke schokken beschikken als medewerknemers in andere beroepen, maar dat ze minder verzekerd zijn tegen permanente schokken. Grotere variabiliteit in de verkoop en de controle door buitenlands kapitaal gaan samen met een lagere verzekering tegen tijdelijke schokken. Een groter risico om failliet te gaan vermindert de mogelijkheden van het bedrijf om haar werknemers inkomensverzekering te bieden. Portugese bedrijven bieden minder verzekering dan hun Italiaanse tegenhangers, hetgeen in overeenstemming is met de hoge loonsflexibiliteit die verbonden is met hun arbeidsmarkt. Wanneer we kijken naar de impact van arbeidsmarktbepalingen, zien we dat in bedrijven met verhoudingsgewijs meer werknemers die het minimumloon verdienen, de inkomens minder gevoelig zijn voor permanente fluctuaties in de vraag. Ook collectieve arbeidsovereenkomsten beperken het vermogen van het bedrijf om idiosyncratische fluctuaties in de vraag om te zetten in inkomensveranderingen.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

355. I.S. SCHINDELE, Three essays on venture capital contracting.
356. C.M. VAN DER HEIDE, An economic analysis of nature policy.
357. Y. HU, Essays on labour economics: Empirical studies on wage differentials across categories of working hours, employment contracts, gender and cohorts.
358. S. LONGHI, Open regional labour markets and socio-economic developments: Studies on adjustment and spatial interaction.
359. K.J. BENIERS, The quality of political decision making: Information and motivation.
360. R.J.A. LAEVEN, Essays on risk measures and stochastic dependence: With applications to insurance and finance.
361. N. VAN HOREN, Economic effects of financial integration for developing countries.
362. J.J.A. KAMPHORST, Networks and learning.
363. E. PORRAS MUSALEM, Inventory theory in practice: Joint replenishments and spare parts control.
364. M. ABREU, Spatial determinants of economic growth and technology diffusion.
365. S.M. BAJDECHI-RAITA, The risk of investment in human capital.
366. A.P.C. VAN DER PLOEG, Stochastic volatility and the pricing of financial derivatives.
367. R. VAN DER KRUK, Hedonic valuation of Dutch Wetlands.
368. P. WRASAI, Agency problems in political decision making.
369. B.K. BIERUT, Essays on the making and implementation of monetary policy decisions.
370. E. REUBEN, Fairness in the lab: The effects of norm enforcement in economic decisions.
371. G.J.M. LINDERS, Intangible barriers to trade: The impact of institutions, culture, and distance on patterns of trade.
372. A. HOPFENSITZ, The role of affect in reciprocity and risk taking: Experimental studies of economic behavior.
373. R.A. SPARROW, Health, education and economic crisis: Protecting the poor in Indonesia.
374. M.J. KOETSE, Determinants of investment behaviour: Methods and applications of meta-analysis.
375. G. MÜLLER, On the role of personality traits and social skills in adult economic attainment.
376. E.H.B. FEIJEN, The influence of powerful firms on financial markets.
377. J.W. GROSSER, Voting in the laboratory.
378. M.R.E. BRONS, Meta-analytical studies in transport economics: Methodology and

applications.

379. L.F. HOOGERHEIDE, Essays on neural network sampling methods and instrumental variables.
380. M. DE GRAAF-ZIJL, Economic and social consequences of temporary employment.
381. O.A.C. VAN HEMERT, Dynamic investor decisions.
382. Z. ŠAŠOVOVÁ, Liking and disliking: The dynamic effects of social networks during a large-scale information system implementation.
383. P. RODENBURG, The construction of instruments for measuring unemployment.
384. M.J. VAN DER LEIJ, The economics of networks: Theory and empirics.
385. R. VAN DER NOLL, Essays on internet and information economics.
386. V. PANCHENKO; Nonparametric methods in economics and finance: dependence, causality and prediction.
387. C.A.S.P. SÁ, Higher education choice in The Netherlands: The economics of where to go.
388. J. DELFGAAUW, Wonderful and woeful work: Incentives, selection, turnover, and workers' motivation.
389. G. DEBREZION, Railway impacts on real estate prices.
390. A.V. HARDIYANTO, Time series studies on Indonesian rupiah/USD rate 1995 – 2005.
391. M.I.S.H. MUNANDAR, Essays on economic integration.
392. K.G. BERDEN, On technology, uncertainty and economic growth.
393. G. VAN DE KUILEN, The economic measurement of psychological risk attitudes.
394. E.A. MOOI, Inter-organizational cooperation, conflict, and change.
395. A. LLENA NOZAL, On the dynamics of health, work and socioeconomic status.
396. P.D.E. DINDO, Bounded rationality and heterogeneity in economic dynamic models.
397. D.F. SCHRAGER, Essays on asset liability modeling.
398. R. HUANG, Three essays on the effects of banking regulations.
399. C.M. VAN MOURIK, Globalisation and the role of financial accounting information in Japan.
400. S.M.S.N. MAXIMIANO, Essays in organizational economics.
401. W. JANSSENS, Social capital and cooperation: An impact evaluation of a women's empowerment programme in rural India.
402. J. VAN DER SLUIS, Successful entrepreneurship and human capital.
403. S. DOMINGUEZ MARTINEZ, Decision making with asymmetric information.
404. H. SUNARTO, Understanding the role of bank relationships, relationship marketing, and organizational learning in the performance of people's credit bank.