Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions

ISBN 978 90 5170 921 6
Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul
This book is no. 431 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.

Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions

Essays over Arbeidsmarkten: Werknemer-Werkgever Dynamiek, Beroepssegregatie en Arbeidsplekomstandigheden

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 27 november 2008 om 09:00 uur
door

Ioan Sebastian Buhai

geboren te Cluj-Napoca, Roemenië

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor:	Prof.dr. C.N. Teulings
Overige leden:	Prof.dr. G.J. van den Berg
	Prof.dr. M.C.W. Janssen
	Prof.dr. J.M. Robin

to my parents

Preface

Working towards a doctoral degree entails much more than just writing a dissertation and this preface is an excellent opportunity to mention a few of the less visible by-products. Inter alia, I would like to use this space to acknowledge the individuals and institutions that, with the benefit of hindsight, had a significant contribution to raising my utility during the long PhD process.

Starting or completing this PhD would not have been conceivable without the guidance of my supervisor, Coen Teulings. I recall a meeting I had in summer 2001, in the former building of the Tinbergen Institute Amsterdam, with the then-General Director of the Tinbergen Institute (TI). At that point I had freshly graduated from University College Utrecht with two Bachelors in several sciences and no clear further direction, except for a fervent interest in continuing the academic road in pretty much any of the disciplines I had been studying. At the end of that half-an-hour "interview", in which we talked about such things as Becker's human capital theory and Ito's Lemma– with Coen managing to fill the whiteboard with graphs during the brief meeting– I knew that I was going to continue with an MPhil and PhD in Economics, at TI, and that I wanted Coen as supervisor. Coen's contagious enthusiasm and joy in doing or explaining Economics research remained throughout all subsequent years a major cause of my appreciation for him. There are additional reasons I am grateful to Coen for. He has been encouraging and supportive throughout, while remarkably patient when needed. I have also learnt tremendously from and enjoyed a lot working with Coen on a couple of exciting joint research projects.

There are a number of individuals who had a direct role in getting this PhD thesis done and in taking the organization around its defence to a final stage. I first want to thank my other co-authors, Elena Cottini, Marco van der Leij, Miguel Portela, Aico van Vuuren and Niels Westergaard-Nielsen, for a gratifying research experience, sharing their knowledge with me and agreeing to have our joint research included in my thesis. Special thanks go to my inner PhD committee, composed of Gerard van den Berg, Maarten Janssen, Jean Marc Robin and my supervisor, for taking the time to evaluate and approve the thesis for public defence, and to Robert Dur and Otto Swank for accepting to be further members of the defence plenary committee. Aico and Marco did a splendid job in translating my dissertation summary in Dutch. Finally, my competent “paranimfen”, Miguel and Robert Sparrow, made sure I would not have to worry about anything concerning the organization of the social events around the defence.

This dissertation has been written during several years, in more than one place: first at the Tinbergen Institute Amsterdam (TIA) and occasionally at the Erasmus University Rotterdam (EUR), next while visiting University College London (UCL), and lately at the Aarhus School of Business (ASB), University of Aarhus. I thank these institutions for providing me with excellent financial and logistic support and with an outstanding academic environment. Part of the expenses concerning my research visits at UCL and ASB were generously funded by grants from the Netherlands Organization for Scientific Research (NWO) and respectively, the Trustfonds of the Erasmus University Rotterdam (VT EUR). Many people at the institutions mentioned above and at other places contributed in various ways to enhancing my happiness, inter alia turning the lonely life of the average PhD into quite the opposite. Ultimately, "PhD-ing" has been indeed one of the most enjoyable activities I have ever done.

The years I spent at the Tinbergen Institute are truly unforgettable. In fact, “those were the days”: the unique mixture between study and research, the first contact with the essence of the academe, the excitement and naiveté of it all. The people I met early on at TI were joyful, relaxed, and really out of the ordinary PhD world in terms of social skills. This subset included such smart, nerdy & super fun types as Miguel, Maria, Robert, Aljaz, Carla, Emily, Ernesto, Gerrit, Hugo, Jasper, Jens, Jurjen, Lev, Marco, Marcos, Martijn, Matthijs, Mauro, Simonetta, Suncica and Wendy. In time, many young & restless PhDs from newer cohorts joined us in ever expanding TI social ventures, the most promising recruits being Ana, Alex, Antonio, Debora, Eva, Marius, Razvan, Romy, Ronald, Rute, Sandra M., Sandra P., Stefan, Tibor, Vali and YinYen. Periodically we also had nice and interesting people visiting TI, such as Marisa, Viktoria, Annette, Laura and Ott, who invariably became enthusiastic friends and colleagues. The members of the TI staff over the years have typically been very agreeable and responsive: I would like to thank Arianne, Carien, Ine, Marian and especially Nora. At TI I was also fortunate to learn a lot, through enlightening courses, seminars, conversations or spot-on advice, from several sharp and friendly Labour Economists: Aico, Bas, Coen, Edwin, Erik, Gerard, Hessel, Jaap, Joop, Maarten, Pieter and Rob. Lastly, I have a few special contributions to mention here. Carla and Miguel have virtually been my adoptive family throughout our stay in Amsterdam; in addition to freely benefiting from their advice in many matters of life, we have attended together countless quality concerts, movies, and parties, and I was privileged to be a

frequent guest at their exquisite Portuguese home dinners. I am indebted to Miguel and Aico for providing guidance and advice for much of my research training, beyond our joint project. I express my thanks to Maria, Miguel and Robert for the numerous politically correct and incorrect talks and loads of fun we had, and I have to single out Robert's brilliant initiative concerning the (in)famous Triple Tours. Ana and Ernesto have been wonderful friends, and excelled as hosts several times. Jens, Marco and Marisa had the dedication and patience to teach me much about Economics topics that I had initially little knowledge about, through exciting joint research projects and many captivating discussions. I owe Ana and Marius many Belgian beers for promptly forwarding to me any mail from TIA, occasionally providing accommodation in Amsterdam and helping with other practical things and information, after my moving to Denmark. I have terrific memories from the time spent in the TI Student Council, having such ideal teammates as Ernesto, Marco, Marcos, Maria, Sandra M., Sandra P., Vali and Wendy.

In the spring of 2005 I was offered the opportunity to visit the Department of Economics at University College London. I wish to express my gratitude to Richard Blundell for inviting me and for being an inspiring and motivating advisor, in matters related both to research and to the academe in general, throughout my stay at UCL. The academic atmosphere I found there was beyond compare and it had a permanent and irreversible effect on my idea about what academe should look and feel like. Though academically rewarding in many respects, my stay in London would not have been nearly half as enjoyable without my interacting with many fine people there, some of whom became great friends and in some cases current or expected future co-authors of projects that will surely revolutionize Economics. This merry crowd of economists included Biagio, Daniele, Hans-Martin, Mario, Erik, Esteban, Josep, Juergen, Katrien, Matthias and Roland. In particular, I ought to credit Biagio for his more than fair share in our joint research on the "Economics of the London Nightlife", acknowledge Daniele's intriguing insights in our critical discussions on worthy vs. non-worthy Economics research & all related matters of life, and thank Hans-Martin and Mario for always interesting conversations and all the great time.

For about the latest three years my host has been the Department of Economics at the Aarhus School of Business, University of Aarhus. I am grateful to Niels Westergaard-Nielsen for offering me the opportunity to come to Aarhus, first as a visiting researcher very much in love with the Danish data (love still lasting), and later as member of the Department and the Center for Corporate Performance at ASB, and for his being friendly, supportive and fun ever since the time we met. At both the ASB Department of Economics and at "the other Department of Economics" of Aarhus University, I have found many

other superb colleagues and friends. A non-exhaustive list of individuals who supplied their fair share as part of a very pleasant social and professional environment necessarily includes Jesper, Frederic, Valérie, Tor, Paul, Mariola, Elena, Chiara, Elisabetta, Erdal, Ija, Jingkun, John, Juanna, Kasia, Lena, Louise, Alex, Anders, Astrid, Camilla, Dale, Henning, Johan, Julia, Lene, Long Hwa, Nabanita, Nicolai, Peter, Roger, Takao and Vibeke. I wish to emphasize a number of further individual contributions to my general welfare. Jesper has remained a great friend and expected terrific co-author, despite my hitherto free-riding on his willingness to help with my frequent moving in Aarhus. Valérie and Frederic have been ideal friends and colleagues ever since we met, and my trusted partners in any social activity in Aarhus or for that matter, in the many conferences around the world attended together. Frederic has also been kind enough to share recently with me two of his interesting master courses, which constitute my first serious teaching experience and which, to my surprise, I discovered to be very enjoyable. From Tor I have often had good advice on many matters, next to plenty of exciting discussions, lots of fun and typically excellent wine to fuel all that. Mariola has been a gentle friend and a perfect host several times. I am grateful to Elena for all the nice chats and for her patience and bearing with my moods. Without Paul and Jesper I would have not understood much about the Danish data. Peter is to be thanked for keeping me happily employed for a great while, despite my apparent desire of being a PhD candidate forever, whereas Lene has been of much help with regards to many cumbersome administrative matters. I am also grateful to Dale and Henning for my inclusion and all further interaction and activities within the Labour Market Dynamics and Growth project, and to John for many recent discussions and initiatives shared.

Throughout my very numerous conferences, summer schools and workshops attended, I befriended a number of nice and talented fellow economists. I particularly want to thank Anna, Carlos, Dimitris, Francois, Julien, Ilan and Sabrina for the good time and stimulating conversations.

Writing this PhD dissertation in Economics could have been done a lot faster, but would have been a lot less enjoyable, without certain non-economist friends in the picture. Firstly, I have to mention my dearest friends, Chrisje, Daniel and Folkert, with whom I pondered any vital and trivial matters that come to mind, and who were and will be there for me, whenever needed, with the often necessary criticism and sometimes required sanity check. I am grateful to them for putting up with my Economics ramblings and for never failing to remind me that there is more to life than cost-benefit analysis and incentives. Ania and Sae, better halves of admirable people, have been gentle friends and excellent hosts several times. I have memories of intriguing long exchanges with Joop, a great friend

with an original approach to life. With Dan I had extensive and exciting discussions in the recent few years and have started collaborating seriously on various ideas of great potential. I need to thank George for many thought-provoking arguments, and for his providing free and unrivalled accommodation conditions whenever I needed to come back to Amsterdam. Although he is truly the laziest when it comes to keeping in touch, I am grateful to Stefano for being my expert guide in our fabulous trips throughout Italy and for all the stirring chats we had. During recent years I did not manage to see much of, but kept in touch and benefited greatly from my enduring friendship with Adina, Ana-Maria, Bahman, Brandon, Camelia, Carmen, Cipri, Colijn, Delia, Flaviu, Iulia, Marcel, Max, Mihai R., Mihai S., Nicu, Thomas and Titus. This review can never be complete without thanking the forever united group of good old Transylvanian friends from Cluj-Napoca, my birth city, who were and remain essential for recharging my batteries every time I return home: Bogdan D., Bogdan F., Gabi, Mircea and Sandu.

More than anything else, I would like to express my infinite gratitude to my beloved parents and sister for their boundless trust, support, and patience.

Sebi Buhai
Aarhus, November 2008

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Structure

This thesis represents an eclectic combination of essays on various dimensions of labour markets. Its four main chapters (2 to 5) should be regarded as independent, self-contained studies. Together they span a wide area of what is typically assigned study object of *modern labour economics*; as the title of the thesis already indicates, the focus will be on the dynamics of the employment relationship between workers and firms, on the arousal and persistence of occupational segregation, income and employment differentials, and respectively, on the link between the firms' work health and safety environment and their financial performance. Each of the papers contributes to its research topic with original and innovative theoretical modelling and/or empirical analysis, often borrowed from other economics fields. This collection of essays embraces for instance approaches as diverse as real options and stochastic calculus, social networks and game theoretical analysis, or microeconometrics techniques at the research frontier within empirical labor economics and empirical industrial organization, applied to five distinct longitudinal or cross-sectional microdatasets[1].

The book proceeds as follows. In Chapter 2, based on joint work with Coen Teulings, we provide brand-new theoretical and empirical insights to an old, but still much debated issue, the tenure profile in wages, highlighting in particular the hitherto neglected but

[1]The datasets used in this thesis are the Panel Study of Income Dynamics (PSID) in Chapter 2; the longitudinal matched employer-employee datasets from Denmark (IDA) and respectively Portugal (Quadros de Pessoal) in Chapter 3; again IDA, plus a longitudinal Danish dataset on firm business accounts (REGNSKAB) and a cross-sectional dataset on workplace health and safety conditions (VOV) in Chapter 5. For the econometric analyses, as well as part of the mathematical computations, the research herein makes use of the following software packages: Maple, Mathematica, Ox, and Stata.

crucial role of the selection on the outside productivity of the worker. Chapter 3, coauthored with Coen Teulings, Miguel Portela and Aico van Vuuren, investigates from a novel perspective the perpetual dilemma of otherwise identical co-workers being paid different wages, emphasizing on both theoretical and empirical grounds the worker's position in the firm's distribution of employee tenures as critical for her bargaining power with regards to job separation probability and wage determination. Both Chapters 2 and 3 are modelling the worker-firm labor market under uncertainty, where the stochastic processes at play are generated by Brownian motions. Hence, next to their research focus, these two essays are connected also by their reliance on stochastic calculus and the application of the real options methodology. The next chapter is in collaboration with Marco van der Leij and provides a creative game theoretical analysis based on social network interactions to a classic economic dilemma: the persistence of wage discrepancies and occupational segregation between groups divided along gender, race or ethnical origin. In Chapter 5, joint with Elena Cottini and Niels Westergård-Nielsen, we present the first empirical research evaluating the relative impact of workplace health and safety conditions on the firm's productivity and mean wage. Finally, the concluding chapter sums up and discusses the results obtained in Chapters 2-5, also addressing directions for future research and potential policy implications. Below I provide brief outlines of the main chapters.

1.2 Outline Chapter 2

There has been quite some literature debating the true magnitude of the wage return to a worker's job tenure, starting with such widely cited papers as Altonji and Shakotko (1987), Abraham and Farber (1987) or Topel (1991), and ongoing to date. In here we approach the problem from a different perspective, by considering first the immediate implication of sizable tenure profiles in wages. From a theoretical perspective, very large returns to tenure are problematic. If they indeed existed, the worker and the firm would spoil large gains from trade when separating; why would the worker ever separate if she loses her tenure profile by doing so? Separations are thus more likely to be layoffs, induced by the firm. But then one can imagine the firm renegotiating the wage with the worker, instead of firing her; some models can explain why this renegotiation process might not be fully efficient, but the size of the wage returns to tenure reported in several papers remains nevertheless puzzling. Given these premises, we explicitly investigate whether the empirical evidence is consistent with *efficient separations*, by modelling jointly the evolution of wages and the distribution of job tenures.

Efficient worker-firm bargaining taken as benchmark, job quits and layoffs become observationally equivalent, as in McLaughlin (1991). Our model explains the correlation between individual wage and tenure from the stochastic evolution of both the worker's productivity in the job and her outside option, that is, her productivity in the best alternative job available. Separation occurs when the value of the worker's job productivity falls sufficiently, relative to her outside productivity. Hence, our story is that wages and tenure are correlated because only jobs where the inside productivity evolves favorably compared to the outside productivity survive. There is no such thing as "classical" return to tenure in this model. The evolution of an individual's within-job log wage is reasonably approximated by a random walk with transitory shocks, as previously found by Abowd and Card (1989) and Topel and Ward (1992), and verified in our Panel Study of Income Dynamics sample. Theoretically, since log in- and outside productivity are assumed to follow continuous random walks, log wages are a linear combination of both, implying that they follow Brownian motions as well. Hence, the difference in the drift between the log wage in the job and the log outside productivity is what we traditionally call "the return to tenure". Starting a job requires an irreversible specific investment, lost upon separation. This combination of investment irreversibility and productivity following a Brownian implies that we can apply the real options methodology, see e.g. Bentolila and Bertola (1990) or Dixit and Pindyck (1994). The predicted hazard rates of our random productivity growth model turn out to be a good fit of the empirical distribution of job exits, our model sharing from this perspective similarities with Mortensen's (1988) dual "on-the-job-training and matching" model or with Jovanovic's (1979b) Bayesian learning model, although it is based on a different theoretical workhorse.

From the distribution of job tenures we estimate, by maximum likelihood, the surplus of the job's productivity above its reservation value, and a drift of this surplus up to a normalizing constant, i.e. the variance of the random walk. The estimated drift is positive, indicating that some jobs will only end by retirement. We use these parameters to compute the evolution of the expected surplus in both completed and incomplete job spells, which will enable us to estimate its impact on wages. Our strategy here is to condition the expected wage growth on both the current and the remaining tenure in a job spell; we can calculate a closed form expression for this expectation. We first show that this expression does not depend on the drift surplus, implying that the evolution of wages in completed spells is completely uninformative on the return to tenure. This is startling, considering all the previous studies that tried to identify the return to tenure precisely from this type of data. Second, we demonstrate that our model can explain the observed concavity in the tenure profile. Since the "true tenure profile", the drift in the difference between inside and outside productivity, is linear by assumption, this concavity is fully due to selection. We are able to test that the estimated variance of the innovation in wages is sufficiently large for selectivity to generate the observed degree of concavity in the tenure profile. Finally, we show that the problem in estimating the tenure profile in wages is not so much the selectivity in the productivity at the job, but rather in the outside productivity. Workers switch jobs only when the outside productivity is high; this effect can be identified from the wage change for job movers. Surprisingly, selectivity in the outside wage turns out to be empirically crucial, accounting for 95% of the tenure profile, though this estimate is somewhat sensitive to misspecification of the model. The estimated "naive tenure profile" is on the high end of the spectrum, almost 3% per year, but almost all of that takes the form of a declining outside productivity instead of a rising inside productivity. As a caveat, our parsimonious framework needs to be adjusted for downward wage rigidity, in order to perfectly fit the data.

1.3 Outline Chapter 3

A large number of studies have investigated why equally productive workers, employed in the same firm and performing the same job, often receive different wages or leave the firm at different times. In this chapter we provide a novel and thorough treatment on this important matter. Using longitudinal matched firm-worker data for Denmark and Portugal, we show that a separation order implied by a *Last-In-First-Out=LIFO rule* (the last worker hired is the first worker fired) is consistent with data, and that there are *returns to seniority* in wages, where seniority is defined as the rank in the tenure hierarchy of the firm's employees, relative to a worker's co-workers. We develop a simple economic theory of why firms and workers would agree on applying a LIFO layoff rule and why that leads to a seniority profile in wages. LIFO is a way to protect the interests of incumbent insiders when hiring and training new workers. Without this protection, the incumbents would have an incentive not to train new workers. LIFO provides protection against layoff for senior workers and hence gives them additional power to bargain for a higher wage, implying seniority profiles in wages. To the extent that this return to seniority is a compensation for the worker bearing part of the cost of specific investment in the employment relationship, the LIFO rule can be interpreted as a protection of the worker's "property right" on her specific human capital in the firm. We show that worker turnover is maximal and the expected job duration is minimal when the surplus and the cost of the specific investments are shared between the worker and the firm in the same proportions, an application of the Hosios (1990) condition.

Our theory is based on a dynamic model of the firm with stochastic product demand and irreversible specific investments for each newly hired worker, similar to Bentolila and Bertola (1990). They calculate optimal hiring and firing points, considering the expected discounted marginal revenue of hiring an additional worker and accounting for the expected moment when it is efficient to fire that worker, by taking as given all workers currently employed by the firm and disregarding any workers that might be hired in the future. Thus, the hiring and firing of each worker can be considered separately, transforming a firm level model into a model of an individual worker, as in Dixit (1989). This turns out to be equivalent to applying a LIFO separation rule. Whereas Bentolila and Bertola (1990) and Dixit (1989) take wages as exogenous, we allow for wage bargaining over the surplus generated by the specific investment at the start of the job spell. To that aim, we apply an idea originated in Kuhn (1988) and Kuhn and Robert (1989). They start from the distinction in trade union theory between the right-to-manage model, where the union bargains for wages and the firm reduces its labor demand in response to the resulting above-market wages, and the efficient bargaining model, where the union and the firm bargain simultaneously over wages and employment, so that employment remains at its efficient level. Kuhn and Robert observe that there is an alternative way for workers to extract rents from the firm, while retaining both right-to-manage and efficiency in employment setting features. Their idea is to bargain for a layoff order and for a wage schedule where inframarginal workers get higher wages than marginal workers. The firm cannot fire expensive inframarginal workers without firing first newly hired workers. When this wage schedule is properly set, the firm will pick the efficient employment level. As a consequence, equally productive workers receive different wages, based on their position in the layoff order. While Kuhn and Robert elaborate their ideas in a static framework, we introduce them in the dynamic model of Bentolila and Bertola, leading to a return to seniority in wages.

In the empirical part of the chapter, we first show that seniority is an important determinant of job separation: junior workers have a larger separation probability than senior workers. This effect comes on top of the duration dependence of the hazard. Next we show that there is a wage return to seniority. We point out that the typical econometric problems in the estimation of the linear return to job tenure are absent in the estimation of the return to seniority, since seniority, unlike tenure, is not perfectly correlated with experience. Seniority increases for example because new workers enter the firm. From that perspective, changes in seniority are correlated with changes in firm size, since an increase in firm size requires new workers to be hired and thus the seniority of the incumbents to increase. Fortunately, in practice seniority is never perfectly correlated with firm size, which would mean that returns to seniority could not be disentangled from the firm size wage effect; seniority also increases by more senior workers leaving the firm, for example due to retirement. In our regressions we use within job spell variation and include both tenure and firm size as controls. We find wage returns to seniority of 1% to 2 % in Portugal, and returns half that range in Denmark. Including seniority reduces the coefficients for tenure and firm size by 5-30 %, suggesting that tenure and firm size served at least partly as proxies for seniority in previous regressions. The return to seniority turns out to be of the same order of magnitude for males and females, but much larger for high- than for low- educated workers, indicating that our estimated results are only lower bounds for the wage returns to "true seniority", with tenure hierarchies constructed for the most appropriate groups of co-workers within the firm.

1.4 Outline Chapter 4

Sociologists and other social scientists often claim that labour economists tend to lose sight of the complexity of individual employment decisions, in the quest for tractable and portable models of worker-firm employment dynamics. One such critique used to be that standard labour economics models fail to account for the importance of social networks in the employer-employee matching process. This view held that a great share of the jobs seem to be obtained through personal connections of the employee and/or employer, next to jobs filled formally. Fortunately this particular critique is no longer relevant since the seminal paper of Montgomery (1991), who has modelled for the first time the role of networks in the labor market. Since then a fast growing literature has been tackling various labour market phenomena, incorporating job contact networks. None of the existing studies has considered however such an investigation with regards to occupational segregation and wage/ unemployment differentials between social subgroups, such as those divided along gender, race or ethnical origin. This is very relevant for several reasons. For one, most studies investigating the extent and determinants of occupational segregation conclude that 'classical' theories such as taste or statistical discrimination cannot alone explain occupational disparities and their remarkable persistence; although several alternative models have been suggested, sizable unexplained parts of the wage differential or combination of occupational segregation and wage and employment disparities could not be fully accounted for. In this light, some researchers emphasized social network interactions as promising avenue for further research in such a context, e.g. Arrow (1998).

Second, there is a lot of empirical evidence across several social science fields– sociology, social psychology and economics, with predilection– that people tend to form connections with others similar to them, what is called *homophily*, *inbreeding bias* or *assortative mating*. Such positive homophily is particularly documented to be robust and sizable for the two gender, racial, and respectively for ethnical subgroups. This evidence should invite consideration of theoretical mechanisms in which individuals within a particular social subgroup would associate with others in the same group more likely than with outsiders, aiming to explain labour market disparities in a subsequent equilibrium analysis. This is what we set up to do in Chapter 4. Our paper builds somewhat on the segregation framework of Benabou (1993), although the two models differ crucially in focus or assumptions of the main mechanisms, and deliver markedly different implications with regards to the social welfare analysis, as detailed in the chapter.

We thus construct a four-stage model of occupational segregation between two homogeneous, mutually exclusive social groups, "Greens" and "Reds", acting in a two-job labor market. In the first stage each individual chooses one of two specialized educations, in order to become a worker. In the second stage individuals randomly form “friendship” ties with other individuals, with a tendency to form relatively more ties with members of the same social group. In the third stage workers use their networks of friendship contacts to search for jobs. In the final stage workers earn a wage and spend their income on a single consumption good. We firstly show, by means of standard arguments on network externalities, that with any small amount of inbreeding bias, a complete polarization in terms of occupations across the two groups arises as a stable equilibrium outcome. We then extend the basic model allowing for “good” and “bad” jobs, in order to analyze equilibrium wage and unemployment inequality between the two social groups; with large differences in job attraction, i.e. wage levels, the main outcome of the model is that one social group "fully specializes" in the good job, while the other group "mixes" over the two jobs. We finally investigate whether society benefits from an integration policy, in that labor inequality between the social groups would be attenuated. To this aim, we analyze a social planner’s first and second-best policy choices. We obtain that segregation is the preferred outcome in the first-best analysis. Assuming that our parameters used for calibration are realistic, a second best social welfare analysis is found to support a laissez-faire policy, where society also becomes segregated, shaped by individual incentives. Hence, overall employment is higher under segregation, while laissez-faire inequality remains sufficiently constrained so that segregation is ultimately an overall socially optimal policy; our social welfare analysis points out relevant policy issues typically ignored in debates concerning anti-segregation legislature.

1.5 Outline Chapter 5

Despite all the attention organizational health and safety measures have received in macro-policy debates in most industrialized countries, it is not a priori obvious which dimensions of the work health and safety environment should be targeted by firms considering to improve them, and in which way an improvement in these conditions would impact corporate productivity and employee welfare. Should one pay equal attention to perceived physical workplace problems such as "noise" or "heavy lifting" burdens or "internal climate conditions", and, respectively, to perceived problems in the psycho-social realm (e.g. decision lattitude of the employees, stress, working with colleagues etc.)? Are these workplace environment dimensions equally relevant in enhancing firm productivity and/or should they be equally compensated for by higher wages when unsolved? The empirical economics literature has been hitherto completely silent on whether better workplace environment –and if so, precisely which dimensions of the "workplace environment"–leads to better firm productivity, and, has not provided either much evidence on whether workplaces where work environment is perceived more hazardous are more likely to pay employees a job hazard premium. In this essay we investigate which firm characteristics associate with good work environment practice and have a first shot at assessing the impact of workplace health and safety conditions on firm performance indicators. While there have been no previous studies focusing on the same topic, this research bears some resemblance to the large and growing recent literature analysing the impact of the firms' industrial resource management system and general reorganization on financial performance, see for instance the papers enumerated in the chapter's introduction.

We link detailed occupational health and safety indicators data from a representative Danish cross-sectional survey of establishments, to the longitudinal register matched employer-employee data, merged with information on the firms' business accounts; the three datasets are source-independent and are merged on a common firm identifier. The first part of our empirical analysis is purely descriptive and consists in estimating binary outcome models of general and specific work environment quality indicators on several employee aggregate characteristics, as well as on proxies of work environment practice, such as having "written work environment rules" or offering "work environment training courses" for all employees. This emphasizes what firm features are associated with good work environment outcomes, in the spirit of Osterman (1994), who looked at the association between firm characteristics and human resource reorganization. The second, and main, part of our analysis consists in estimating standard Cobb-Douglas production functions, augmented with employees' aggregated characteristics such as the proportion of females, proportion of unskilled workers, average human capital in the firm, and the specific work environment indicators that were used as dependent variables in the previous logit analyses. The longitudinal dimension of the register firm data enables us to estimate these augmented production functions in two stages, using either fixed firm effects estimation or system-generalized method of moments à la Blundell and Bond (1998, 2000) in the first stage, where we only use the production inputs and aggregate employees'characteristics, and respectively OLS of the mean (over time) residuals resulting from the first stage on the cross-sectional work environment indicators, in the second stage. This closely follows the strategy by Black and Lynch (2001) for evaluating the impact of firm re-organization on firm productivity, allowing us to address eventual endogeneity biases due to unobserved time-invariant firm effects, as well as simultaneity of classical inputs and output in the production function. We also investigate the explanatory power of work environment conditions and other employee aggregate characteristics in accounting for between-firm mean wage differentials, with firm fixed effects estimation in a first stage, and a second stage that uses the time-average residual from the first stage, regressed on the workplace condition indicators. A significant improvement relative to Black and Lynch (2001) is that in our dataset we observe all firm characteristics over time, not just the evolution of the firm standard production inputs, and in particular we can also proxy for likely time-variant unobservables such as managerial ability, which might otherwise remain correlated with the work condition indicators in the second stage estimation, by instrumenting for changes and lagged levels of the proportion of managerial positions over time.

The following factors are found to have explanatory power in accounting for the variation in the workplace conditions among firms: the proportion of managerial positions, all-employee work environment courses offered in the firm and, to less extent, the proportion of female employees in the firm's workforce and prioritizing work environment practice at the firm. In terms of effects of work environment indicators on firm performance, our results suggest that, among all specific workplace conditions considered, only improvement in some of the hardcore physical dimensions of workplace environment, "internal climate" and respectively, "repetitive and strenuous work activity", impacts the firm aggregate productivity, and that this impact is sizable. At the same time, the only workplace health and safety condition with explanatory power in the between-firm mean wage differential is the same "internal climate", suggesting a compensating wage differential story.

Chapter 2

Tenure Profiles and Efficient Separation in a Stochastic Productivity Model

2.1 Introduction

A large empirical literature has looked at wage returns to job seniority, using a whole arsenal of econometric techniques, see Farber (1999) for a survey. The conclusions of this research still diverge, despite analyzing data from the same countries (mainly the USA) or even the same longitudinal datasets (mostly the PSID): while some authors find that large estimated returns are spurious and wage returns to tenure are actually very small, e.g. Altonji and Shakotko (1987), Abraham and Farber (1987), Altonji and Williams (1997, 2005), Abowd et al (1999), others confirm large and significant wage returns close to cross-section estimates, e.g. Topel (1991), Dustmann and Meghir (2005), Buchinsky et al (2005). Here we provide a new direction for investigating the wage-tenure relationship. From a theoretical point of view, large "true" returns to tenure are problematic. Were there really large returns, the worker-firm match would spoil large gains from trade at the moment of separation. Why would a worker separate when he loses his tenure profile by doing so? Hence, separation is likely to be induced by the firm, what we call a layoff.

[0] This chapter is based on the revised version of Buhai and Teulings (2006), recently resubmitted to the *Review of Economic Studies*. We thank two anonymous referees, our editor Bernard Salanié, Rob Alessie, Gadi Barlevi, Richard Blundell, Jeff Campbell, Bas van der Klaauw, Jim Heckman, Dale Mortensen, Derek Neal, Randall Olsen, Miguel Portela, Robert Shimer, Chris Taber, Aico van Vuuren, Robert Waldmann, and participants in seminars at the Univ. of Chicago, Northwestern Univ., Federal Reserve Bank of Chicago, London School of Economics, Univ. College London, Univ. of Copenhagen, Tinbergen Institute, and several conferences and workshops in Europe and US, for helpful comments. We also thank Nick Williams and Lennart Janssens for assistance in handling the data. The usual disclaimers apply.

But why would the worker and the firm not renegotiate the wage instead of separating? Although some models, such as efficiency wage models, can explain why this renegotiation process might not be fully efficient, the size of the wage returns to seniority reported in some papers remains puzzling. In fact, the empirical evidence offers support for at least some form of renegotiation. For instance Jacobson, LaLonde and Sullivan (1993) have shown that displaced workers face severe wage cuts of up to 25% just before separation. This paper addresses explicitly whether the existing evidence is consistent with efficient separations by modelling simultaneously the evolution of wages and the distribution of job tenures.

We take efficient bargaining as benchmark. Hence, quits and job layoffs are observationally equivalent, as in McLaughlin (1991). The model explains the correlation between wages and job tenure from the random evolution of both the job's productivity and the outside option. Separation occurs when the value of the productivity in the job falls sufficiently compared to the productivity of the outside option. This outside option is the productivity in the best alternative job that is available at that point in time. We refer to productivity at the job and in the best alternative as the *inside* and *outside productivity* respectively. By some form of bargaining, wages at the job are a linear combination of the inside and the outside productivity. Then, wages and tenure are correlated because only jobs where inside productivity evolves favorably compared to the outside productivity survive. Hence, there is no such thing as "the" return to tenure in this model. In some jobs wages go up because the job's productivity value evolves favorably. In others wages go down for *mutatis mutandis* the same reason. However, the latter group is gradually eliminated from the stock of ongoing employment relations just because there are no options for mutually gainful renegotiation left and hence separation becomes efficient.

The evolution of an individual's within-job log wage is reasonably described by a random walk with transitory shocks, as previously found by Abowd and Card (1989), Topel (1991) and Topel and Ward (1992). We verify that hypothesis in our PSID sample. Whereas this observation received little attention among labor economists, we take it as cornerstone of our modelling. Both log in- and outside productivity are assumed to follow a random walk. Our model implies that log wages are a linear combination of both, which implies that log wages in the job follow a random walk as well. Hence, the difference in the drift between the log wage in the job and the log outside productivity is what we traditionally call "the return to tenure".

Starting a job requires an irreversible specific investment, which is lost upon separation. Hence, this investment has an option value. The combination of irreversibility and productivity following a random walk implies that we can apply the theory of real options, see for example Dixit (1989), Bentolila and Bertola (1990) and Dixit and Pindyck (1994), compare Teulings and van der Ende (2000). The predicted hazard rates of this model are well in line with the empirical distribution of the job exits. Our model is similar to Mortensen's (1988) dual "on-the-job-training and matching" model. While we focus on firm tenure, our model could equally well be applied to industry or occupation tenure, as suggested by Neal (1995).

From the distribution of job tenures we are able to estimate the surplus of the job's productivity above its reservation value and a (linear) drift of this surplus, up to a normalizing constant (the variance of the random walk). We obtain a positive drift surplus, indicating that some 10% of all jobs will end only by retirement. We use these parameters to compute the evolution of the expected surplus in both complete and incomplete job spells, which will enable us to estimate its impact on wages. The typical problem in this literature is that the researcher observes the outside productivity only at job start and at job separation, assuming that the worker has a new job immediately afterwards. At job start, the worker chooses the best alternative that is available at that moment, which is by definition equal to the outside productivity. Our estimation procedure exploits both pieces of information on the outside productivity. To that end, we apply an idea first explored by Abraham and Farber (1987), conditioning the expected wage growth on both the current and the remaining tenure at that job. We can calculate a closed form expression for this expectation. As a first result, we show that this expression does not depend on the drift surplus. This implies that the evolution of wages in completed spells is uninformative on the return to tenure. This is a remarkable conclusion given the fact that so many papers have tried to identify the return to tenure from this type of data. The only sources of information on the return to tenure are the distribution of completed tenures and the evolution of wages in incomplete job spells. The fat right tail in the tenure distribution, with many jobs never ending, is an indication of large returns to tenure: the return to tenure is so high that separation is rarely efficient.

Secondly, we show that our model can explain the observed concavity in the tenure profile. Since the "true" tenure profile, the drift in the difference between inside and outside productivity, is linear by assumption, this concavity is fully due to selection. One could argue that our identification procedure relies heavily on functional form assumptions. However, there is one strong test of our assumptions: the estimated variance of the innovation in wages is sufficiently large for selectivity to generate the observed degree of concavity in the tenure profile.

Thirdly, we show that the problem in estimating the tenure profile in wages is not so much the selectivity in the inside productivity (and hence in the wage rate at the job), but in the outside productivity. Workers switch jobs only when the outside productivity is high. This source of selectivity usually receives less attention than the selectivity in the inside productivity. We show that this effect can be identified from the wage change for job movers. Surprisingly, selectivity in the outside wage turns out to be an empirically important phenomenon, accounting for 95% of the tenure profile, though this estimate is sensitive to misspecification of the model. In particular, our estimation results provide some indications of downward rigidity in wages, as discussed for example by Beaudry and DiNardo (1991), who find that within a job spell wages go up in the upturn, but do not go down in the downturn. However, in our estimation results, this gap is filled by an additional wage decline for job changers. This downward rigidity does not fit the efficient bargaining hypothesis. The estimated tenure profile is on the high end of the spectrum, almost 3% per year, though almost all of that takes the form of a declining outside productivity instead of a rising inside productivity. If we were to exclude this part of the profile, our estimates would be on the low end of the spectrum, 0.15% per year.

The paper is structured as follows: the model is discussed in Section 2, the empirical analysis in Section 3 and Section 4 concludes.

2.2 The random productivity growth model

2.2.1 Model assumption

Consider a labor market in continuous time, where both workers and firms are risk neutral. We focus on a single cohort of homogeneous workers. We normalize our measure of time t such that it is also equal to the workers' experience. There is no disutility of effort, so that the workers' utility depends on their expected lifetime income only. Each firm offers a single job, of which the productivity P_t evolves according to a geometric Brownian with drift; P_t is job specific. At the moment a worker is hired for a vacant job, a specific investment has to be made which is partly paid by the firm and partly by the worker and which is irreversibly lost upon separation between the worker and the job. However, the firm retains the property right on the vacant job. Hence, the firm can hire a new worker for that job at any future time, provided that the cost of the specific investment is paid again. This cost of the specific investment is verifiable. There is no search cost involved from either party in finding a new job: an unemployed worker can just pick the most attractive vacancy that is available at that time, at zero cost. Since there are always vacant jobs available, a worker has a shadow price R_t, which is equal to the return in the best alternative vacant job, net of the cost of investment for that job. For the sake of convenience, we treat this shadow price as an exogenous variable here. Like P_t, it evolves according to a geometric Brownian with drift; since workers are homogeneous, R_t is common to all of them. Both workers and firms are perfectly informed about the current value of the P_t's for each job and of R_t, but their future evolution is unknown. The value of the specific investments for a job starting at time t is $R_t I$. One can think of I as the cost of investment measured in units of labor time and of R_t as the price of one unit at time t. Using lower cases to denote the logs of the corresponding upper cases, the law of motion of p_t and r_t, for $t > s$, is characterized by a bivariate normal distribution:

$$\begin{bmatrix} p_t - p_s \\ r_t - r_s \end{bmatrix} \sim N\left[(t-s)\underline{\mu}, (t-s)\Sigma\right]$$

where:

$$\Sigma = \begin{bmatrix} \sigma_p^2 & \sigma_{pr} \\ \sigma_{pr} & \sigma_r^2 \end{bmatrix}, \underline{\mu} = \begin{bmatrix} \mu_p \\ \mu_r \end{bmatrix} \tag{2.1}$$

The worker and the firm bargain over the surplus of the productivity of the job above the shadow price of a worker, $P_t - R_t$. This bargaining is efficient: as long as there is a surplus, the worker and the firm will agree on a sharing rule.

2.2.2 Value of a job and a vacancy

Three assumptions made above greatly simplify the analysis. (i) The risk neutrality of both players implies that the allocation of risk is irrelevant; only expected values matter. (ii) The verifiability of investment implies that there are no hold up problems: the distribution of future surpluses $P_t - R_t$, $t > s$, is irrelevant for the timing of the investment decision, since the cost of the specific investment $R_s I$ can always be shared between the worker and the firm according to their relative bargaining power. Hence, the investment decision will maximize the joint expected surplus of the worker and the firm. (iii) Efficient bargaining implies that separation decisions will also maximize the joint expected surplus. Hence, separation occurs at mutual consent when there are no gains from trade left. Quits and layoffs are therefore observationally identical, as in McLaughlin (1991). For the sake of convenience, we shall refer to separations as the firm firing the worker, though they can be both quits or layoffs. Given these assumptions, wage setting and separation decisions can be analyzed separately, since, in the spirit of the Coase theorem, hiring and firing decisions maximize the joint expected surplus, regardless of its precise distribution.

First, we analyze hiring and firing. Since hiring is in fact an irreversible investment, while firing is an irreversible disinvestment, both can be analysed using real option theory, see Dixit and Pindyck (1994). The easiest way to analyze this problem is to assume that workers always get paid their shadow price R_t. Then, hiring and firing simply maximize the expected value of the firm. Let $V(p_t, r_t)$ and $J(p_t, r_t)$ be the expected present value of a vacancy and respectively of a job, as functions of p_t and r_t. Applying Ito's lemma, the Bellman equations for both value functions read, compare Dixit and Pindyck (1994: pp.140-141):

$$\begin{aligned} \rho J &= \exp(p_t) - \exp(r_t) + \mu_p J_p + \mu_r J_r + \frac{1}{2}\sigma_p^2 J_{pp} + \sigma_{pr} J_{pr} + \frac{1}{2}\sigma_r^2 J_{rr} \\ \rho V &= \mu_p V_p + \mu_r V_r + \frac{1}{2}\sigma_p^2 V_{pp} + \sigma_{pr} V_{pr} + \frac{1}{2}\sigma_r^2 V_{rr} \end{aligned} \tag{2.2}$$

where we leave out the arguments of $J(\cdot)$ and $V(\cdot)$ for convenience and where ρ denotes the interest rate. The term $\exp(p_t) - \exp(r_t)$ in the first equation is the value of current output minus the wage of the worker; the other terms capture the wealth effects due to changes in the state variables p_t and r_t: the first order derivatives capture the effect of the drift in both state variables, the second order derivatives capture the effect of their variance. For optimal hiring and firing, value matching and smooth pasting conditions should be satisfied:

$$\begin{aligned} J(p_S, r_S) &= V(p_S, r_S) + \exp(r_S) I \\ V(p_T, r_T) &= J(p_T, r_T) \\ J_p(p_S, r_S) &= V_p(p_S, r_S) \\ J_r(p_S, r_S) &= V_r(p_S, r_S) + \exp(r_S) I \\ V_p(p_T, r_T) &= J_p(p_T, r_T) \\ V_r(p_T, r_T) &= J_r(p_T, r_T) \end{aligned} \tag{2.3}$$

where S is the moment of hiring and T is the moment of firing. The first two conditions are the value matching conditions for hiring and firing respectively, which state that the values before and after the hiring or firing should be equal. The first condition for hiring states that at the moment of hiring S the value of a job must be equal to the value of the vacancy plus the cost of specific investment. The second condition for firing states that at the moment of firing T the value of the job must be equal to the value of a vacancy. The last four conditions are the smooth pasting conditions. Value matching conditions impose value equality before and after hiring and firing; on top of that, smooth pasting conditions require that slight variations in the stochastic variables p_t and r_t should not affect the value equality, since hiring and firing decisions are irreversible. Hence, a decision maker should not regret her decision a minute later, due to slight variations in p_t or r_t. Smooth pasting requires thus the partial derivatives of the value matching condition with respect to p_t and r_t to be zero. These conditions and the Bellman equations (2.2) jointly determine $J(\cdot)$ and $V(\cdot)$.

Define $b_t \equiv p_t - r_t$; b_t is the log of the relative surplus of P_t over R_t. By (2.1), we have:

$$\begin{aligned} b_t - b_s &\sim N\left[(t-s)\mu, (t-s)\sigma^2\right] \\ \sigma^2 &\equiv \sigma_p^2 + \sigma_r^2 - 2\sigma_{pr} \\ \mu &\equiv \mu_p - \mu_r \end{aligned} \tag{2.4}$$

Since p_t is specific for each job, so is b_t.

Proposition 1 *The value functions $J(\cdot)$ and $V(\cdot)$ can be written as:*

$$\begin{aligned} J(p_t, r_t) &= \exp(r_t)\, j(p_t - r_t) \\ V(p_t, r_t) &= \exp(r_t)\, v(p_t - r_t) \end{aligned} \tag{2.5}$$

where $j(\cdot)$ and $v(\cdot)$ satisfy:

$$\begin{aligned} \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) j &= \exp(b_t) - 1 + \left(\mu + \sigma_{pr} - \sigma_r^2\right) j' + \frac{1}{2}\sigma^2 j'' \\ \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) v &= \left(\mu + \sigma_{pr} - \sigma_r^2\right) v' + \frac{1}{2}\sigma^2 v'' \end{aligned}$$

where we leave out the argument of $j(\cdot)$ and $v(\cdot)$ for convenience. The value matching *and smooth pasting conditions at the moment of job start and job separation read:*

$$
\begin{aligned}
j\left(b^S\right) &= v\left(b^S\right)+I \\
v\left(b^T\right) &= j\left(b^T\right) \\
j'(b^S) &= v'(b^S) \\
v'(b^T) &= j'(b^T)
\end{aligned}
$$

where b^S, b^T are the values of b_t at the moment of hiring and firing respectively.

Proof: The proposition follows directly from substitution[1] of equation (2.5) in the Bellman equations (2.2) and the value matching and smooth pasting conditions (3.6).■

The smooth pasting conditions for p_t and r_t are identical, so we are left with only two independent smooth pasting conditions. The factor $\rho-\mu_r-\frac{1}{2}\sigma_r^2$ is a modified discount rate, which accounts for the fact that future revenues are discounted at a rate ρ, but increase in expectation at a rate $\mu_r+\frac{1}{2}\sigma_r^2$ due to the drift and the variance of R_t. The hiring and separation rules depend therefore purely on b_t: a vacancy should be filled at the first time t that $b_t=b^S$, a worker should be fired from the job at the first time t that $b_t=b^T$. This proposition characterizes the decision problem of the firm by two second order differential equations, four boundary conditions and two decision parameters, b^S and b^T. This is exactly the "basic model" described by Dixit and Pindyck (1994; ch. 5.1-5.2), to whom we refer for the subsequent arguments. The solution to the two differential equations has four constants of integration. Two of these constants have to be zero due to a transversality condition. The constants of integration reflect the option value for the firm of hiring and firing a worker. The option value of hiring converges to zero when $b_t\to 0$, while the option value of firing converges to zero when $b_t\to\infty$. These constraints can only be satisfied by setting two constants of integration equal to zero. Hence, the four boundary conditions determine four unknown parameters: b^S, b^T, and the two remaining

[1] We use:

$$
\begin{aligned}
J_p &= \exp\left(r_t\right)j', J_{pp}=\exp\left(r_t\right)j'' \\
J_r &= \exp\left(r_t\right)\left(j-j'\right), J_{rr}=\exp\left(r_t\right)\left(j-2j'+j''\right) \\
J_{pr} &= \exp\left(r_t\right)\left(j'-j''\right)
\end{aligned}
$$

and likewise for $V(\cdot)$.

constants of integration. One can prove $b^T < 0 < b^S$. Hiring occurs at the first moment that b_t rises to $b^S > 0$. Hence, $P_t > R_t$, because the firm has to recoup the cost of investment and because the investment is irreversible, so that the firm loses the option value of hiring later, while in the meantime b_t might fall below b^S again. Subsequent firing occurs at the first moment that b_t falls to b^T. Hence, $P_t < R_t$ because the firm accepts some losses before firing the worker, since it loses the option value of firing the worker later.

2.2.3 Job tenure distribution

The next step is to analyze the distribution of job tenure in a job spell. The duration of a job spell is a stochastic variable, equal to the time it takes the random variable b_t to travel down from b^S to b^T. Analogously to the probit model, where the variance of the error term is non-identified because we observe only whether the indicator variable is positive or negative, the standard deviation of b_t is unidentified in this model because, for any time t, we observe only whether the spell is still incomplete, implying $b_t - b^T > 0$ ever since the start of the job spell. We can therefore normalize all parameters by σ. For each job spell, we define $\tau \equiv t - S, \tau \geq 0$, and $\Theta \equiv T - S, \Theta > 0$; τ is the incomplete tenure, while Θ is the completed tenure of that job spell. Define:

$$\begin{aligned} \Omega_\tau &\equiv \frac{b_t - b^T}{\sigma} \\ \Omega &\equiv \frac{b^S - b^T}{\sigma} > 0 \\ \pi &\equiv \frac{\mu}{\sigma} \end{aligned}$$

Thus Ω_τ is a Brownian with drift π and unit variance per unit time. By construction $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$. Θ is determined by the time it takes Ω_τ to pass the barrier $\Omega_\tau = 0$ for the first time. This process satisfies the "First Passage Time" distribution, which has been applied previously by Lancaster (1972) for modelling strike durations, and by Whitmore (1979) for job spells. The unconditional density of $\Omega_\tau = \omega$ reads:

$$\frac{1}{\sqrt{\tau}}\phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}}\right)$$

where $\phi(\cdot)$ is the standard normal PDF. However, a job spell is completed if and only if Ω_ζ has not been negative for any $\zeta \in [0, \tau]$. Hence, we are interested in the density of Ω_τ conditional on $\Omega_\zeta > 0, \forall \zeta \in [0, \tau]$. For this conditioning, we can apply the Reflection Principle, first discussed by Feller (1968) for the case without drift, $\pi = 0$: there is a one-to-one correspondence between trajectories of Ω_τ from Ω to ω which have crossed the barrier $\Omega_\tau = 0$ at least once, and trajectories of Ω_τ from $-\Omega$ to ω. These trajectories should therefore be subtracted to obtain the conditional density of Ω_τ. Define: $g(\omega, \tau) \equiv \Pr(\Omega_\tau = \omega \wedge \Theta > \tau)$. It satisfies, see e.g. Kijima (2003, p.185-187)[2]:

$$g(\omega, \tau) = \frac{1}{\sqrt{\tau}} \left[\phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}} \right) - e^{-2\Omega\pi} \phi\left(\frac{\omega + \Omega - \pi\tau}{\sqrt{\tau}} \right) \right] \tag{2.6}$$

where $\phi(.)$ is the standard normal density function. The first term in square brackets is the unconditional density; the second term is the effect of the conditioning. By the Reflection Principle, the latter is the density of trajectories of Ω_τ from $-\Omega$ to ω. The factor $e^{-2\Omega\pi}$ corrects for the differential effect of the drift on the density for upward and downward trajectories. By integrating out ω we obtain the cumulative distribution of jobs surviving at τ, $\overline{F}(\tau) = \Pr(\Theta > \tau)$:

$$\overline{F}(\tau) \equiv \Phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}} \right) - e^{-2\Omega\pi} \Phi\left(\frac{-\Omega + \pi\tau}{\sqrt{\tau}} \right) \tag{2.7}$$

where $\Phi(.)$ is the standard normal CDF. This expression above is identical to Whitmore (1979: eq. 2). The distribution of Θ is therefore fully specified by two parameters, the initial distance from the separation threshold Ω and the drift π. Hence, Ω and π can be identified from data on job tenures, while the parameter σ cannot. The corresponding density function is minus the derivative of $\overline{F}(\tau)$ with respect to τ:

$$f(\tau) = \frac{\Omega}{\tau\sqrt{\tau}} \phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}} \right) \tag{2.8}$$

[2]Kijima (2003, pp. 185-187) derives the precise expression of the transition density for our case, namely for a standard Brownian with drift $\pi > 0$, starting at $\Omega_0 = \Omega > 0$, and one absorbing barrier at $\Omega_\Theta = 0$. Many other books on stochastic processes derive the similar conditional density but for a standard driftless Brownian $\pi = 0$ and/or starting at $\Omega_0 = 0$ and/or with positive absorbing barrier $\Omega_\Theta = a > 0$. See for instance Cox and Miller (1965, pp. 219-223), Feller (1968, vol.2, p. 328), Zhang (1998, p. 208-218) etc. As shown by these authors, one can use various methods to derive the expression, the Reflection Principle being the most intuitive.

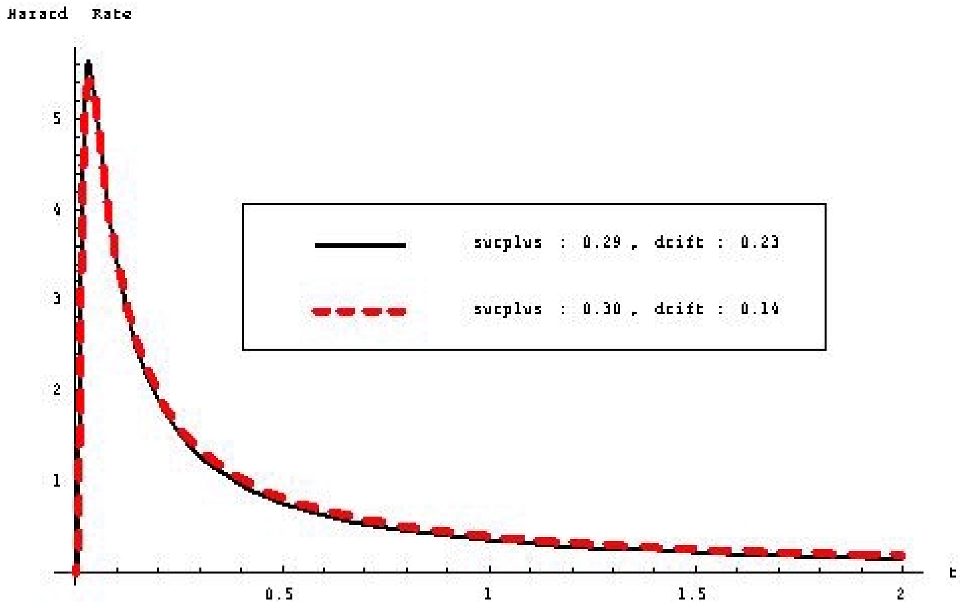

Figure 2.1: Predicted job hazards

where we have used $\phi\left(\frac{\Omega+\pi\tau}{\sqrt{\tau}}\right) = e^{-2\Omega\pi}\phi\left(\frac{-\Omega+\pi\tau}{\sqrt{\tau}}\right)$. The job exit rate is then given by $f(\tau)/\overline{F}(\tau)$. It is straightforward to check that the exit rate is hump shaped, starting from 0, reaching a peak at τ^*, $0 < \tau^* < \frac{2}{3}\Omega^2$, and afterwards either declining monotonically to 0 when the drift is positive $\pi > 0$ or to ${}^1/{}_2\pi^2$ when the drift is negative $\pi < 0$. Farber (1994), Teulings and Van der Ende (2000) and Horowitz and Lee (2002) have documented this hump shaped pattern using NLSY data. A positive drift implies a non exhaustive behavior, where some jobs never end. The fraction of surviving job spells for $\pi > 0$ is given by the value of the survivor function (2.7) for $\tau \to \infty$, hence by $1 - e^{-2\Omega\pi}$. In Figure 2.1, we plot the exit rates for $\Omega = e^{-1.20} \simeq 0.30, \pi = 0.14$ and respectively for $\Omega = e^{-1.24} \simeq 0.29, \pi = 0.23$, the mean estimated values for Ω and π, see Section 3, Table 2 below. In both cases the peak is reached at $\tau \simeq 0.04$ years. Since $\pi > 0$, the hazard rate converges to zero and a positive fraction of the jobs, about 10 %, will never end.

2.2.4 Tenure profile in wages

Sharing rule of surpluses and wages

We extend the model with an explicit sharing rule of surpluses during the course of the job spell. Ideally, we would derive this sharing rule from an explicit bargaining game, such as Nash bargaining. For the sake of convenience, we use however a simpler approach, by imposing the log linearity of the sharing rule *a priori*, and deriving the intercept of that rule from the assumption of efficient bargaining.[3] According to this rule, the worker's log wage w_t satisfies:

$$w_t = r_t + \beta \left(b_t - b^T\right) + \psi = r_t + \overline{\sigma}\Omega_\tau + \psi \tag{2.9}$$

where $\overline{\sigma} \equiv \beta\sigma$. The parameter β can be interpreted as the worker's bargaining power. If $\beta = 0$, the worker receives a wage proportional to her shadow price R_t, while if $\beta = 1$, she receives a wage proportional to her productivity at the job, P_t. To close the model, the parameter ψ and the worker's share in the cost of investment remain to be determined. Since we do not need their expressions for the subsequent empirical analysis, we only provide an heuristic argument here. Let $Q(p_t, r_t)$ be the worker's asset value of holding a job, net of the discounted expected value of her shadow price R_t. Analogous to Proposition 1, $Q(\cdot)$ can be written as:

$$\begin{aligned} Q(p_t, r_t) &= \exp(r_t)\, q(p_t - r_t) \\ \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) q &= \exp\left[\beta\left(b_t - b^T\right) + \psi\right] - 1 + \left(\mu + \sigma_{pr} - \sigma_r^2\right) q' + \frac{1}{2}\sigma^2 q'' \end{aligned}$$

leaving out the argument of $q(\cdot)$ in the second line. The first term in the second line captures the wage at the job, compare equation (2.9), while the second term captures the outside wage. Note that we have divided both sides of the equation by $\exp(r_t)$. Efficient bargaining implies that it is optimal for the worker to quit when $b_t = b^T$. Hence, value matching and smooth pasting conditions must apply: $q\left(b^T\right) = 0, q'\left(b^T\right) = 0$. The value matching condition states that at the moment of separation the net asset value of continuation is zero. Again, the solution of the differential equation has two constants of integration, one of which has to be zero due to a transversality condition. Hence, the solution to the differential equation and the two boundary conditions determine the constant of integration and ψ. Finally, the assumption of verifiability of the cost of specific investment at the moment of job start implies that the worker's share in this cost must be equal to the net value of holding a job at the moment of job start, $\exp(r_S)\, q\left(b^S\right)$. Hence, this share is equal to $q\left(b^S\right)/I$.

[3] Nash bargaining would lead to a linear, instead of a log linear, sharing rule. Apart from this, the two approaches are identical. Note that Nash bargaining satisfies the assumption of efficient bargaining.

Selectivity in tenure profiles

Equation (2.9) implies that log wages within a job follow a Brownian with drift $\mu_r + \overline{\sigma}\pi$; μ_r is the sum of the return to experience for that cohort of workers and the secular growth of real wages due to technological progress common to all cohorts; $\overline{\sigma}\pi$ is the deterministic part of the tenure profile. Were the realizations of Ω_τ independent of the completed job tenure Θ, $\overline{\sigma}\pi$ could be estimated easily. Empirically, wages are observed in discrete time. Hence, for estimation, the easiest way would be to first difference equation (2.9):

$$\begin{aligned} \text{job stayers} &: \quad \Delta \mathrm{E}\left(w_t | 1 < t - S < \Theta\right) = \mu_r + \overline{\sigma}\pi \\ \text{job changers} &: \quad \Delta \mathrm{E}\left(w_t^* | t = T\right) = \mu_r - \overline{\sigma}\pi\left(\Theta - 1\right) \end{aligned}$$

where Δ is the first difference operator and where the superscript $*$ indicates that we compare log wages in the new and the old job; hence, Δw_T^* compares the starting wage in the new job to the wage one year before separation in the old job. Hence, $\overline{\sigma}\pi$ can be estimated from data on job changers. However, in completed job spells, Ω_τ is correlated to Θ for three reasons: (i) $\Omega_0 = \Omega$, (ii) $\Omega_\Theta = 0$, and (iii) $\Omega_\zeta > 0, \forall \zeta, 0 \leq \zeta < \Theta$. For the sake of brevity, we refer to this information set as $A\left(\Theta\right)$. Our strategy for estimation is to calculate $\mathrm{E}(\Omega_\tau | A\left(\Theta\right))$ and to enter its first difference as a regressor in a regression on Δw_t's within completed job spells. *Mutatis mutandis,* the same applies to incomplete spells. Let Ψ be the incomplete tenure at the last date for which data are available. Again, there are three pieces of information: (i) $\Omega_0 = \Omega$, (ii) $\Theta > \Psi > \tau$, and hence (iii) $\Omega_\zeta > 0, \forall \zeta, 0 \leq \zeta \leq \Psi$. We refer to this second information set as $B\left(\Psi\right)$. Again, we can calculate $\mathrm{E}(\Omega_\tau | B\left(\Psi\right))$ and use its first difference as a regressor, see also Van der Ende (1997).

Proposition 2 *$E(\Omega_\tau|A(\Theta))$ and its derivatives satisfy:*

$$
\begin{aligned}
E(\Omega_\tau|A(\Theta)) &= 2\sqrt{m(\tau)\tau}\phi\left(\sqrt{m(\tau)/\tau}\Omega\right) - \left(\frac{\tau}{\Omega} + m(\tau)\Omega\right)\left[1 - 2\Phi\left(\sqrt{m(\tau)/\tau}\Omega\right)\right] \\
m(\tau) &\equiv \frac{\Theta - \tau}{\Theta} \\
\lim_{\tau\to 0} \frac{dE(\Omega_\tau|A(\Theta))}{d\tau} &= \frac{1}{\Omega} - \frac{\Omega}{\Theta} \\
\lim_{\tau\to \Theta} \frac{dE(\Omega_\tau|A(\Theta))}{d\tau} &= -\infty \\
\frac{d^2E(\Omega_\tau|A(\Theta))}{d\tau^2} &< 0
\end{aligned}
$$

Proof See Appendix A.■

This proposition implies that $\mathrm{E}(\Omega_\tau|A(\Theta))$ does not depend on the tenure profile in wages, $\overline{\sigma}\pi$. Hence, conditional on the model that we specified, the evolution of wages in completed job spells does not provide any information whatsoever on the tenure profile in wages. Given the many papers that have tried to estimate tenure profiles from data on completed job spells, this is a staggering conclusion. The intuition for this result is that an increase in $\overline{\sigma}\pi$ has two offsetting effects on $\Delta\mathrm{E}(\Omega_\tau|A(\Theta))$. On the one hand, it raises the deterministic part of the tenure profile, so that the change in the unconditional expectation $\Delta\mathrm{E}(\Omega_\tau)$ goes up. On the other hand, it makes separation a less likely event, so that the condition $A(\Theta)$ becomes more selective: the non-deterministic part of $\Delta\Omega_\tau$ must have evolved unfavorably for a job spell to end after Θ, even though the deterministic part of the tenure profile pushes Ω_τ up. This conclusion depends crucially on the assumption of efficient bargaining. This assumption dictates

$$w_T - w_S = r_T - r_S - \overline{\sigma}\Omega$$

see equation (2.9). Hence, irrespective of the steepness of the tenure profile $\overline{\sigma}\pi$ or the length of the job spell Θ, log relative wages decline by $\overline{\sigma}\Omega$ over the duration of the spell. However, as noted in Section 2.3, π can be estimated from the tenure distribution. Efficient bargaining implies that this distribution is informative on the tenure profile, since under efficient bargaining, a higher tenure profile implies that jobs will survive longer. We return to the issue of identification in Section 3.

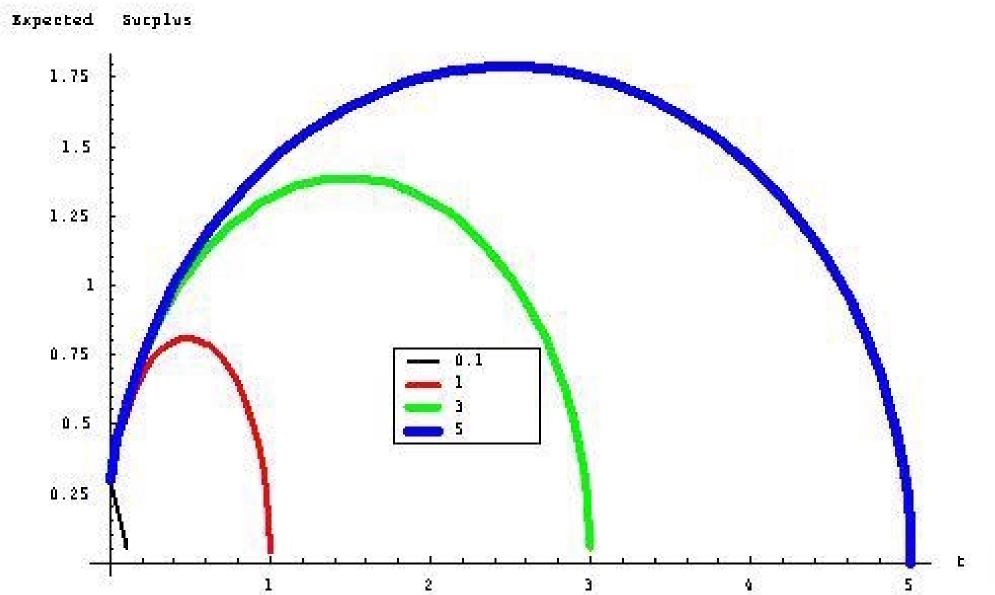

Figure 2.2: Expected surplus in completed job spells

The third line of Proposition 2 says that the initial slope of $\mathrm{E}(\Omega_\tau|A(\Theta))$ is negative for short spells, $\Theta < \Omega^2$, even when the drift is positive, $\pi > 0$. For these spells, $\mathrm{E}(\Omega_\tau|A(\Theta))$ must decline immediately for $\Omega_\Theta = 0$. The fourth line shows that the expected surplus declines infinitely fast just before separation. This result is consistent with empirical evidence by Jacobson, LaLonde and Sullivan (1993) on the decline in relative wages in the period just before firing. The final line shows that the second derivative is always negative. Hence, $\mathrm{E}(\Omega_\tau|A(\Theta))$ is concave in τ; it is monotonically decreasing for short spells $\Theta < \Omega^2$ and it is hump-shaped for longer spells. The tenure profile is plotted for $\Omega = 0.30$ and for various values of Θ in Figure 2.2. For $\Theta \leq 0.1$ years the tenure profile is monotonically decreasing, while for larger Θ it is concave. The top of the profile is increasing in Θ, showing the importance of conditioning on the eventual tenure.

Contrary to the case of completed spells, there is no explicit expression for $\mathrm{E}(\Omega_\tau|B(\Psi))$. Hence, we use numerical integration, see Appendix B. Figure 2.3 presents the trajectory of $\mathrm{E}(\Omega_\tau|B(\Psi))$ for $\Omega = 0.30, \pi = 0.14$ and various values of Ψ. $\mathrm{E}(\Omega_\tau|B(\Psi))$ is increasing in Ψ. The reason is that a higher value of Ψ provides more information on Θ since $\Theta > \Psi$. Hence, higher values of Ψ imply a greater selectivity. Were there no selectivity, then the trajectory would be linear, $\mathrm{E}(\Omega_\tau|B(\Psi)) = \mathrm{E}(\Omega_\tau) = \pi t$. The trajectories are strongly concave, implying that selection plays an important role. This offers an explanation of the observed concavity of tenure profiles in log wages: the underlying profile might be linear and the observed concavity might simply be due to selection. Contrary to the completed

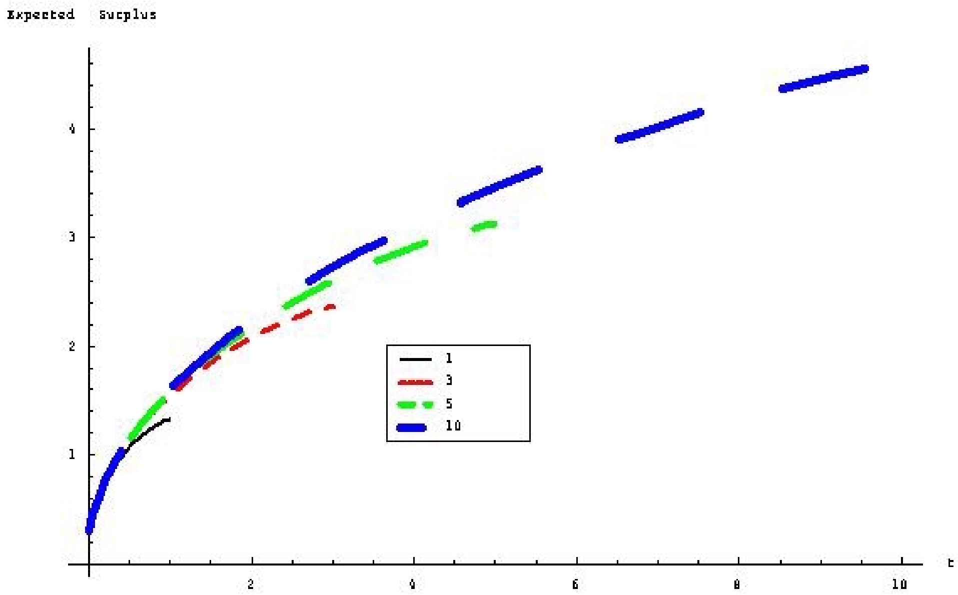

Figure 2.3: Expected surplus in incomplete job spells

spells case, incomplete spells do provide information on the drift π. Nevertheless, the impact of the drift is negligible compared to that of selectivity, as documented by Figure 2.4, which compares the trajectories of $\mathrm{E}(\Omega_\tau | A(\Theta))$, $\mathrm{E}(\Omega_\tau | B(\Psi))$, and $\mathrm{E}(\Omega_\tau)$. The concavity outweighs the linear trajectory by far, at least for the first five years. In Figure 2.5 we plot $\mathrm{E}(\Omega_\tau | A(\Theta))$ and $\mathrm{E}(\Omega_\tau | B(\Psi))$ for long job durations, $\Theta = 10, 20$ and respectively $\Psi = 10, 20$.

Expected within-job and between-job wage growth

We apply the conditional expectations of Ω_τ for the empirical analysis of the growth in w_t and r_t. We observe r_t only at the moment of job change. Hence, whereas we can use information of within job wage growth for the analysis of w_t, we have to rely on between job wage growth for the analysis of r_t. For this purpose, we linearly decompose the random variables $[\Delta p_t, \Delta r_t]$ in two orthogonal components Δb_t and Δz_t, such that $\mathrm{Cor}(\Delta b_t, \Delta z_t) = 0$, and that the effect of Δz_t on Δr_t and Δw_t is equal to unity. Then:

$$\begin{aligned} \Delta r_t &= \Delta z_t - \gamma \beta \Delta b_t = \Delta z_t - \gamma \overline{\sigma} \Delta \Omega_\tau \qquad (2.10) \\ \Delta w_t &= \Delta z_t + (1-\gamma) \beta \Delta b_t = \Delta z_t + (1-\gamma) \overline{\sigma} \Delta \Omega_\tau \end{aligned}$$

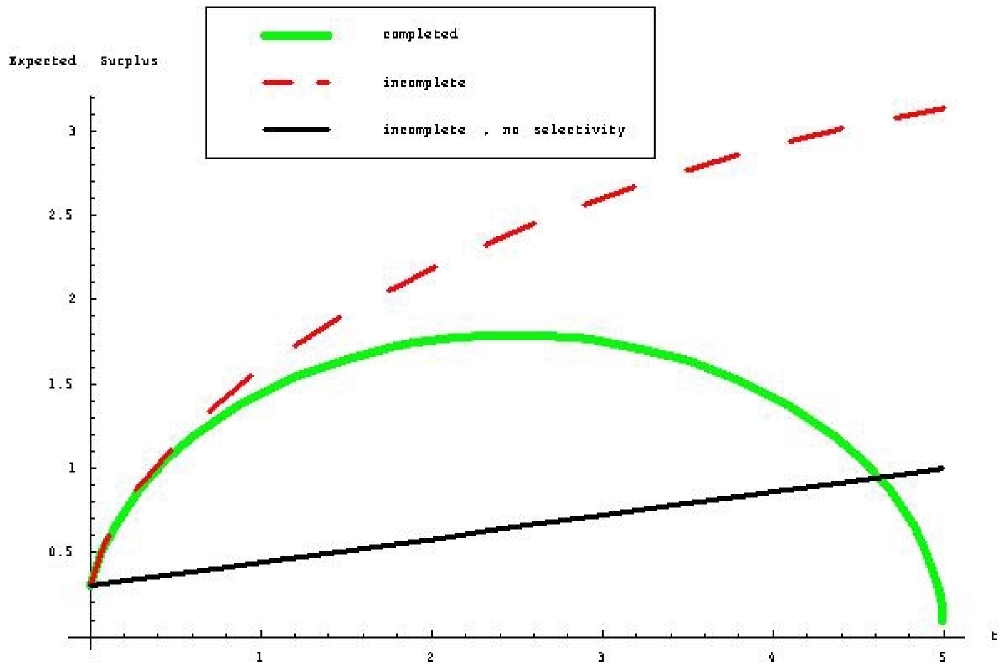

Figure 2.4: Selectivity versus drift in the expected surplus

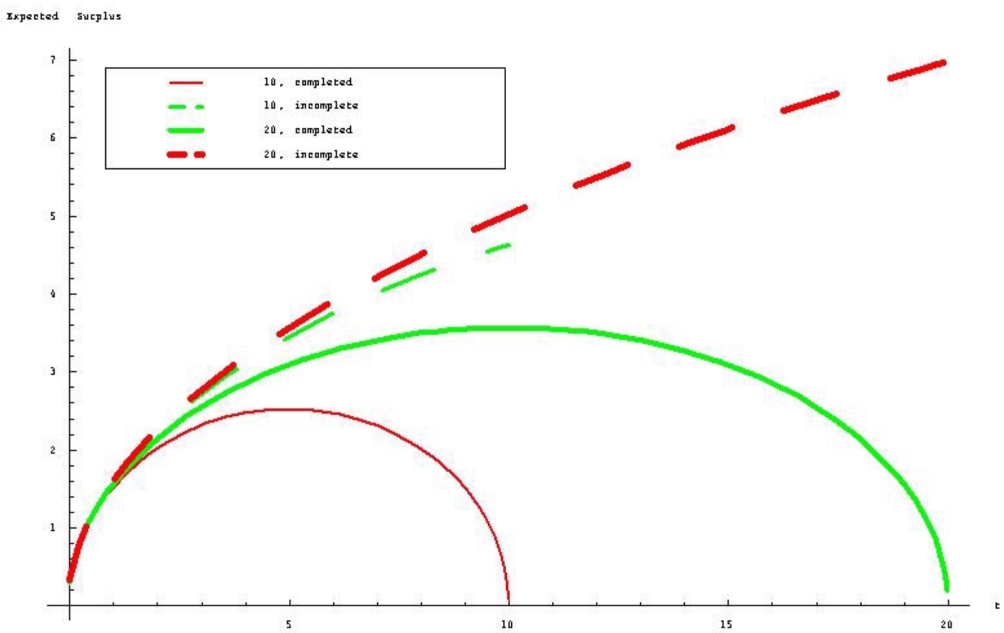

Figure 2.5: Expected surplus in long spells

Given the joint normality of Δp_t and Δr_t, such a decomposition always exists. The advantage of this decomposition is that, since separation decisions are determined by the evolution of b_t and since Δb_t and Δz_t are uncorrelated, selectivity affects Δb_t, but not Δz_t. The parameter γ can be expressed in terms of the covariance matrix Σ and the bargaining power β, but that is of little help here. It is more useful to interpret it as a reflection of the correlation between the match surplus and the reservation wage. In the one extreme case where $\gamma = 0$, we can write $\Delta p_t = \Delta r_t + \Delta b_t$, with both right-hand side variables being uncorrelated. Then Δr_t reflects the evolution of the general human capital of the worker, which evolves independently of the value of the specific capital in the present job, Δb_t. Hence, the duration of the actual job is fully determined by its own (mis)fortune. Though the distinction between quits and layoffs makes little sense in this model, separations look like layoffs in this case: the firm fires the worker since she is no longer productive. In the opposite extreme case where $\gamma = 1$, we can write $\Delta r_t = \Delta p_t - \Delta b_t$, again with both right-hand side variables being uncorrelated. Now Δp_t reflects the evolution of the general human capital of the worker; Δb_t reflects the specific evolution of outside opportunities, e.g. new technologies emerging in other firms. Separations look like quits in this case: the worker quits because she can get a better job elsewhere. In this case, the selectivity of job relocation is not so much that of the type "only good jobs survive outside offers", but more of the type "only good outside offers kill the job".

It is immediately clear from equation (2.10) that there is no hope of identifying β and σ separately from data on wages, since only their product $\overline{\sigma}$ shows up in the final expression for Δw_t. There is a clear intuition for this result. We do not observe the productivity p_t or the surplus b_t, but only the share that goes to workers as a wage payment w_t.[4]

Taking expectations in the second equation of (2.10) yields:

$$\mathrm{E}(\Delta w_t | A(\Theta)) = \mu_z + (1-\gamma)\overline{\sigma}\Delta\mathrm{E}(\Omega_\tau | A(\Theta)) \tag{2.11}$$

where $\mu_z \equiv \mathrm{E}(\Delta z_t)$. This equation applies for completed spells; replacing the condition $A(\Theta)$ by $B(\Psi)$ yields the equation for incomplete spells. As discussed in Section 2.3, Ω and π can be estimated from the distribution of completed tenures. These parameters are sufficient statistics for the calculation of $\mathrm{E}(\Omega_\tau | A(\Theta))$. Equation (2.11) allows the estimation of $(1-\gamma)\overline{\sigma}$ by OLS. However, γ and $\overline{\sigma}$ are not separately identified. The intuition is that we observe the current wage of a worker, but not her shadow price for

[4] Teulings and Van der Ende (2000) work out a method that allows the estimation of σ. They interpret hours spend on training at the start of a job spell as a source of variation in I. The estimated covariation between I and Ω and the assumption of the absence of hold up problems allow the identification of σ.

an alternative job.

For job changers, we can write a similar equation:

$$\mathrm{E}(\Delta w_T^*) = \mu_z + (1-\gamma)\overline{\sigma}\Delta \mathrm{E}(\Omega_\Theta | A(\Theta)) + \overline{\sigma}\Omega \tag{2.12}$$

where, as before, the superscript $*$ indicates that we compare log wages in the new and the old job. $\Delta \mathrm{E}(\Omega_\Theta | A(\Theta))$ is the wage decline in the old job the year before separation.[5] It is always negative, see Figure 2.2. The term $\overline{\sigma}\Omega = \overline{\sigma}(\Omega_0 - \Omega_\Theta)$ reflects the wage increase due to entering the new job. The final term of equation (2.12) allows the separate identification of γ and $\overline{\sigma}$. The intuition is that at the moment a worker changes jobs, the wage jumps up by $\beta\left(b^S - b^T\right) = \overline{\sigma}\Omega$. This conclusion is very much similar to the standard result in this literature that information on job movers is crucial for the estimation of the tenure profile.

Till sofar we focused on the model's implications for the first moment of Δw_t. However, the model has also implications for the second moment of Δw_t. Taking the variance in the second equation of (2.10) yields:

$$\sigma_w^2 = \sigma_z^2 + (1-\gamma)^2\overline{\sigma}^2$$

where $\sigma_w^2 \equiv \mathrm{Var}(\Delta w_t)$ and $\sigma_z^2 \equiv \mathrm{Var}(\Delta z_t)$. Equation (2.9) implies:

$$w_T^* - w_S = r_T - r_S = z_T - z_S - \gamma\overline{\sigma}\Omega \tag{2.13}$$

where, as before, the superscript $*$ indicates that we refer to the starting wage in the new job starting at T, and where we use the first equation of (2.10), and $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$ for the second equality. Since $\mathrm{Var}(\Omega) = 0$ and $\mathrm{Var}(z_T - z_S) = \Theta\mathrm{Var}(\Delta z_t)$ we have:

$$\mathrm{Var}\left(w_T^* - w_S\right) = \Theta\sigma_z^2$$

[5]Implicitly, we assume here that separation takes place exactly at the end of the year of observation. This is an important assumption, since wages decline steeply in the last year before separation, see Figure 2.2. If separation occurs earlier on during the year of observation, part of the fall in wages during the last year before separation is captured by the previous observation.

Hence:

$$(1-\gamma)^2\overline{\sigma}^2 = \sigma_w^2 - \Theta^{-1}\text{Var}\left(w_T^* - w_S\right) \tag{2.14}$$

This test relates the observed variance of wage changes σ_w^2, net of the variance of the overall shock z, to the degree of concavity in wages, $(1-\gamma)\overline{\sigma}\Delta\text{E}(\Omega_\tau|A(\Theta))$, in other words, it tests whether there is sufficient yearly variation in wages to generate the observed concavity from selection.[6]

2.3 Empirical analysis

2.3.1 The data

We use a dataset based on a PSID extract of 18 waves, covering the years 1975 through 1992, same as the one used by Altonji and Williams (1997, 1999). Our model does not work well when employed people consider other alternatives than switching to another job, like retirement, leaving the labor force or taking up full time education. The availability of these other alternatives yields two problems. First, we do not observe the reservation wage at the point of separation when people do not accept another job. Second, with only one alternative to the present job, the decision problem is simply whether a particular indicator switches signs. With more alternatives, that choice process becomes far more complicated. Therefore we restrict the sample to people who do not switch in and out the labor force regularly and for whom retirement is not a relevant option: white male heads of household with more than 12 years of education (we also drop the few observations that have a missing value for education) and less than 60 years of age. Our reasoning is similar to the one used in Mincer and Jovanovic (1981), who also use job separation synonymous to job change, thereby also defining labor mobility as change of employer and excluding other alternatives, which are minor phenomena in the case of the full-time male working force. Furthermore, we restrict the attention to those individuals that were employed, temporarily laid off, or unemployed at the time of the survey, and were not from Alaska or Hawaii. Finally, we discard all observations on unionized jobs.[7]. We use the tenure and experience measures constructed by Altonji and Williams (1999). Table 2.1 presents

[6] This test does not account for effect of the condition $A(\Theta)$ on the observed variance in Δw_t. Due to this conditioning, the observed variance is a lower bound for σ_w^2. The expressions for the conditional variance are extremely cumbersome. However, for longer spells, $\Theta > 5$, the effect of conditioning on the variance is small.

[7] The previous working paper version of this paper includes unionized spells and controls for other covariates, see Buhai and Teulings (2006).

summary statistics of the data. Since we do not need wages in the tenure distribution analysis, observations with missing wage information are included in that analysis. One can distinguish four types of job spells. Apart from the distinction between completed and incomplete spells (right censoring), one can also make a distinction between spells that start before the time span covered by the data, and spells that start afterwards (left censoring). The lower half of the summary statistics table informs on the number of spells for each of these four types.

Table 2.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	Observations
logwage[(1)]	2.42	0.52	0.17	4.82	13660
tenure (years)	6.67	7.42	0.08	43.69	15504
experience (years)	14.58	9.21	0.12	43.69	16179
No. of obs. discarded from the data in Altonji and Williams(1997)					10351

Dataset for estimating the tenure distribution parameters[(2)]	
Number of individuals	2421
Total number job spells	4681
- started before the observation range	1512
- started within the observation range	3169
Completed job spells	1712
- started before the observation range	372
- started within the observation range	1340
Incomplete job spells	2969
- started before the observation range	1140
- started within the observation range	1829

[(1)]reported average hourly wage, deflated using the implicit price deflator with 1982 base year

[(2)]subset of data summarized in the top panel, keeping one observation for each job spell

2.3.2 The parameters of the tenure distribution

Our estimation strategy uses the recursive feature of our model, that Ω and π can be estimated from the tenure distribution, and that these parameters can then be used to calculate $\Delta\mathrm{E}(\Omega_t|A(\Theta))$ and $\Delta\mathrm{E}(\Omega_t|B(\Psi))$, that are entered in the analysis of wage dynamics. Ω and π are estimated by maximum likelihood, using the density function (2.8). For the theoretical analysis, we have treated both parameters as constants that do not depend on worker characteristics. Empirically, one can expect that workers choose their optimal job type according to their characteristics. Hence, Ω and π are likely to differ according to both observed and unobserved worker characteristics. As observed worker characteristics we enter only experience at the start of the job, S. Since we deal with longitudinal data, we can take into account random worker effects. We do not consider random job effects for both theoretical and empirical reasons. From a theoretical point of view, our assumption of a frictionless market for alternative job opportunities, where the only constraint on instantaneous mobility is the specific investment in the present job and not the cost of getting another job offer, each worker type will choose that job type that fits best her comparative advantages, like in Sherwin Rosen's hedonic world of kissing curves. Hence, job characteristics are implied by worker characteristics. From an empirical point of view, we observe each job only once, thus we have no basis for identifying random job effects other than from functional form assumptions. Taking into account that Ω has to be positive, the following specification for Ω and π is adequate:

$$\begin{aligned}\Omega &= \exp\left(\beta_{\Omega 0} + \beta_{\Omega}\widehat{S} + u_{\Omega}\right) \\ \pi &= \beta_{\pi 0} + \beta_{\pi}\widehat{S} + u_{\pi}\end{aligned} \tag{2.15}$$

where u_{Ω} and u_{π} are normally distributed random worker effects with zero mean and standard deviations σ_{Ω} and σ_{π}, and where hat on $\widehat{S}$ denotes deviations from its mean over jobs. Hence, the intercept can be interpreted as the mean value for Ω and π respectively. We assume both random effects to be uncorrelated. Then, the contribution to the log likelihood function for an individual reads:

$$\log L = \ln \int\int \prod_{j=1}^{J} \overline{F}(\Psi_j)^{1-d_j} \cdot f(\Theta_j)^{d_j} d\Phi\left(\frac{u_{\Omega}}{\sigma_{\Omega}}\right) d\Phi\left(\frac{u_{\pi}}{\sigma_{\pi}}\right) \tag{2.16}$$

where j is the j^{th} job held by the worker, where d_j is a dummy variable, taking the value $d_j = 1$ if the job spell is completed and the value $d_j = 0$ otherwise. There are two reasons why we have to make amendments to the simple likelihood function in equation (2.16).

First, we could restrict the estimation to job spells starting within the observation range in the PSID extract. However, this means that we would not consider any of the jobs started before they were first reported in the data. By construction, this would limit the maximum completed tenure in the data to the maximum time span covered by the PSID sample, that is 17 years. Since long tenures contain relevant information, we want to include also spells started before their first wave in the PSID. We know either Θ_j or Ψ_j for these spells and we can compute experience at the beginning of a job by subtracting current tenure from current experience. However, we observe these spells only conditional on the fact that they have lasted till the start of our observation period. We should correct the log likelihood function for this condition:

$$\log L = \ln \int \int \overline{F}(\tau_1)^{-1} \prod_{j=1}^{J} \overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{2.17}$$

where τ_1 is the tenure in the job at the start of its observation in the PSID (for which $j = 1$).

Second, since the PSID collects data at a yearly interval, job spells completed in less than a year are underreported. We know the elapsed tenure in months at the first moment a job spell is observed, by a retrospective question[8], but we do not know whether there has been another job spell between the job observed a year ago and the job observed now. Since the hazard rate implied by our model is hump shaped, with the hump likely to be within the first year, c.f. Farber (1994), this phenomenon is expected to have a large impact on the estimation results. We are likely to overestimate Ω and π, since we miss part of the short tenures in our data. Hence, we have to correct for this form of left censoring. One solution to this problem is to use a similar conditioning as in equation (2.17), where τ_j is the initial tenure at the first moment the job is observed (measured in months in PSID). However, this approach does not use the distribution of these τ_j's itself.[9] We can use this distribution if we are prepared to make the additional assumption that the starting date of job spells is distributed uniformly over the first year. Then, the density $q(\cdot)$ of initial dates of spells that started throughout the year and are still incomplete at the end of the year satisfies:

$$q(\tau) = \frac{\overline{F}(\tau)}{\int_0^1 \overline{F}(x)\, dx}$$

The total contribution to the likelihood of a spell with initial tenure τ and completed tenure Θ is therefore:

$$\frac{f(\Theta)}{\overline{F}(\tau)} q(\tau) = \frac{f(\Theta)}{\int_0^1 \overline{F}(x)\, dx}$$

[8]Initial tenures are either reported or inferred by making them consistent with the latest reported tenures- see Altonji and Williams (1999 and previous working versions).

[9]Maximum likelihood estimation using this approach yields a huge hump in the hazard rate, which implies a much higher share of spells shorter than a year that can be justified from the distribution of τ_j for jobs started after the first wave.

Hence, the log likelihood reads:

$$\log L = \ln \int \int \prod_{j=1}^{J} \frac{\overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j}}{\int_0^1 \overline{F}(x)dx} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{2.18}$$

The log likelihood that accounts both for jobs starting before their first reporting in the PSID and for the left censoring of spells shorter than a year started after the first wave of the PSID, can thus be written as:

$$\log L = \ln \int \int \prod_{j=1}^{J} \frac{\overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j}}{\overline{F}(\tau_1)^{I(j=1)} \left(\int_0^1 \overline{F}(x)dx\right)^{I(j\neq 1)}} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{2.19}$$

where $I(y)$ is the indicator function, taking value 1 if y is true and value 0 otherwise. We report results for (2.18), where we use only the sample of jobs that start within their observation period, and for (2.19), where we use all job spells, including those started before they are first observed in the PSID[10]. The estimation results are presented in Table 2.2.

Theoretically, the results for the two likelihood functions should be identical. The theoretical hazards for both models look indeed almost identical (c.f. Figure 2.1 above), the only difference being the height of the peak, lower for the case where we use all job spells. The same can be concluded also by inspecting Table 2.2, where the estimated intercepts are very similar, while the coefficients for experience at job start are virtually the same, for both Ω and respectively π. The positive effect of experience on the drift π would be consistent with the idea that workers start their career with some initial job hopping, before settling down in a job that fits one's comparative advantages best. Furthermore, the intercept for π is positive and large in both estimations. In both cases, there are hardly observations for which π is negative. This implies that some job spells will last until the retirement of the worker. The fraction of jobs that never end, for mean values of the parameters, is about 10%.

[10]In order to estimate the log-likelihood functions above, we used simulated maximum likelihood, cf. Stern (1997). Sampling from a joint normal distribution with mean 0 and variances σ_π^2 and σ_Ω^2 and using a sampling size of 500 sampling points (the results are robust to altering the sampling dimension to any size between 100 and 500 sampling points) we achieved strong convergence in a reasonable number of iterations. We used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method for convergence of derivatives, allowing for a tolerance of 1E-4 times the absolute value of the log likelihood.

Table 2.2: ML Tenure distribution parameters

	Small sample(1)		Large sample(2)	
Variable	**Drift π**	**Dist Ω**	**Drift π**	**Dist Ω**
Intercept	0.226**	-1.243 **	0.141**	-1.197**
(st. errors)	(0.023)	(0.087)	(0.002)	(0.016)
Initial experience	0.009**	-0.006	0.012**	0.002
(st. errors)	(0.003)	(0.010)	(0.0002)	(0.002)
Random worker effects σ	0.309**	0.002	5.75e-007	3.66e-005
(st. errors)	(0.053)	(1.219)	(0.002)	(0.013)
Observations (job spells)	3169		4681	

(1)Small sample= sample of job spells starting within the range covered in the PSID

(2)Large sample= sample of all job spells

All covariates are taken in deviations from their means over jobs

Significance levels: † : 10% ∗ : 5% ∗∗ : 1%

One remarkable conclusion is that there are no unobserved random worker effects when we use the sample of all jobs spells, while there is unobserved heterogeneity in the drift for the sample including only the shorter spells within the observation range. Since the long spells started before the first wave contain crucial information, we focus on the estimation results obtained from the full sample of job spells in the subsequent wage dynamics analysis.

As a test of the goodness of fit of the model, we compute the predicted distribution of incomplete tenures after 32 years experience and compare that to the observed distribution.[11] Figure 2.6 depicts both the predicted and the empirical density of incomplete job spells.There is a reasonable correspondence between both densities. The peak in the first year is overestimated, but otherwise the shapes of the two densities are identical. Note the small peak in the density for short incomplete spells, which is due to the hump shape pattern in the hazard: if your job ends for instance in the last five years before the end of the observation period, there is a substantial probability that you experience further separations afterwards due to the peak in the hazard rate, leading to a peak of short incomplete tenures. Close alignment of the predicted and the empirical densities suggests that our model works well.

[11]This density is calculated by a recursive scheme. We divide the 32 years time period in 32×256 subperiods. We calculate the distribution of completed tenures for jobs starting at the beginning of the career, in the first subperiod. For some of these jobs, $t > 32$, which is the density of incomplete tenures of 32 years. Then we calculate the distribution of completed tenures for jobs starting in the second subperiod, which is the number of jobs started in the first subperiod that separate in the second. We add this number to the corresponding completed tenures of the jobs started in the first subperiod. Then we calculate the completed tenure for jobs started in the third subperiod, etc. In these calculations we account for the effect of experience at the job start on the parameters Ω and π, which we estimated above.

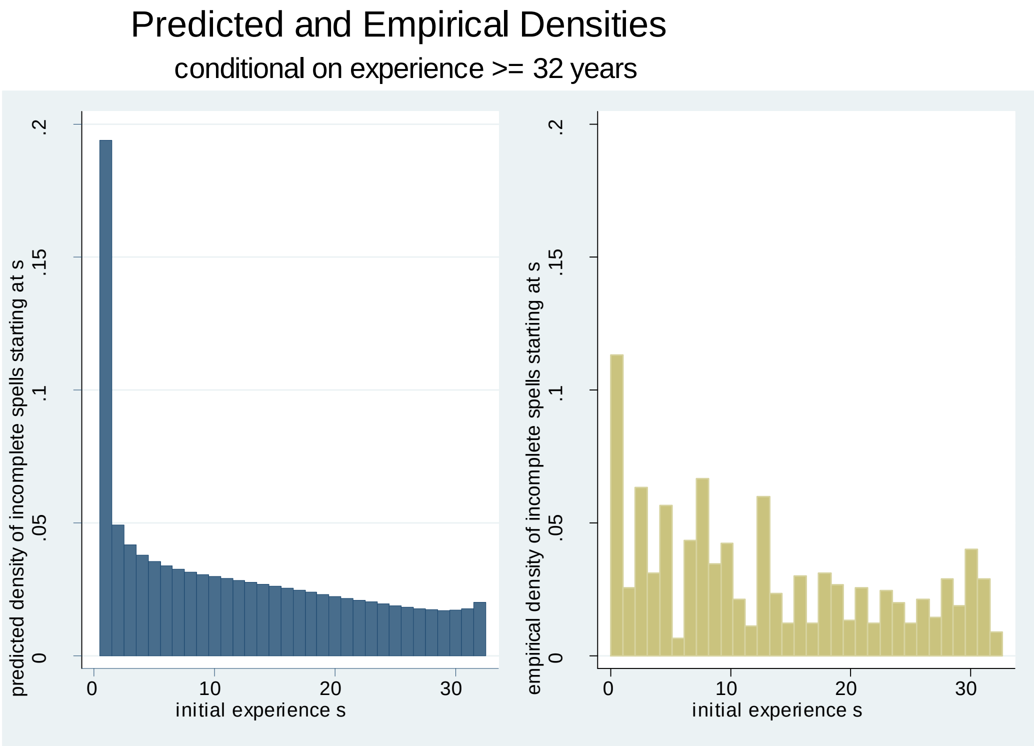

Figure 2.6: Density of incomplete job spells with exit option

2.3.3 Test of the random walk hypothesis in wages

To prepare the ground for our formal analysis of wage dynamics, we document some stylized facts on the stochastic dynamics of wages. In particular, we verify that log wages follow a random walk using the methodology applied by MaCurdy (1982), Abowd and Card (1989) and Topel and Ward (1992). First, we run a regression of within-job log wage differentials on a number of controls:

$$\Delta w_t = \delta_0 + \delta_t t + \delta_\tau \tau + \Delta u_t \tag{2.20}$$

Though our theoretical model allows only for linear tenure and experience profiles, we enter both variables (t=tenure, τ=experience) in the first differenced equation to capture the eventual concavity of both profiles. Tenure is entered as a proxy for the concavity due to the selection process as discussed in Section 2.4. We run separate regressions for job stayers and switchers[12]. The regression results are displayed in Table 2.3.

Table 2.3: Within- and between-jobs wage change regressions

	Within-jobs	**Between-jobs**
Intercept	0.0458**	0.1109**
(std. err.)	(0.0042)	(0.0183)
Tenure	0.0001	-0.0034
(std. err.)	(0.0003)	(0.0028)
Experience	-0.0016**	-0.0039**
(std. err)	(0.0003)	(0.0014)
Observations	8802	1243
SER(1)	0.1951	0.3584
R^2	0.0049	0.0092

Significance levels : † : 10% * : 5% ** : 1%

(1) SER= standard error of the regression (root mean square error)

In the regression for stayers, we find some evidence for concavity in the experience, but not in the tenure profile. For job switchers we find that experience at separation negatively affects the wage differential to the new job, but that tenure at separation has no statistically significant effect. The results concerning experience are consistent with the findings in Buchinsky et al (2005)[13].

[12] In the case of job switchers we regress the change in the log wages from one job to the other, on the last tenure and respectively experience, at the moment of separation.

[13] Buchinsky et al (2005) also find a positive effect of the seniority at job separation on the starting wage in the new job and evidence of concavity in the seniority profile for job stayers, which are both statistically

Next, we construct a covariogram of residuals Δu_t of the within-job wage change regression, see Table 2.4.

Table 2.4: Residual autocovariances for within-job wage innovations

Lag	Autocovariance	Standard Deviation
0	.0380**	.0020
1	-.0130**	.0017
2	-.0012*	.0006
3	.0009	.0007
4	.0004	.0007
5	-.0006	.0007
6	.00004	.0008
7	-.0005	.0008
8	-.0004	.0009
9	.0015	.0010
10	.00007	.0012
11	-.0011	.0015
12	.0004	.0013
13	.0016	.0012
14	-.0017	.0013
15	-.0022	.0017
16	-.0013	.0018

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Residuals are strongly negatively correlated to their first lag, while autocovariances for longer lags are statistically insignificant beyond lag 2. This outcome is similar to results obtained by MaCurdy(1982), Abowd and Card (1989) and Topel and Ward (1992). Our covariogram is thus typical of an MA(2) process or even an MA(1) once we note that the second order lag autocovariance is close to 0. For simplicity, we focus on the MA(1) case. We decompose the stochastic time-variant component of the wage equation from (2.20) in a martingale persistent component e_t and a transitory component η_t:

$$\begin{aligned} u_t &= e_t + \eta_t \\ \Delta e_t &= \varepsilon_t \end{aligned} \tag{2.21}$$

insignificant in our case. This difference can be due to slight differences either in the specification or in the PSID sample they use, as can be seen from comparing summary statistics.

where η_t and ε_t are i.i.d. with $\text{Var}(\eta_t) = \sigma_\eta^2$ and $\text{Var}(\varepsilon_t) = \sigma_w^2$. Then:

$$\begin{aligned} \text{Var}(\Delta u_t) &= \sigma_w^2 + 2\sigma_\eta^2 \\ \text{Cov}(\Delta u_t, \Delta u_{t-1}) &= -\sigma_\eta^2 \\ \text{Cov}(\Delta u_t, \Delta u_{t-k}) &= 0, k > 1 \end{aligned}$$

This is a good description of the pattern of autocovariances in Table 2.4. Hence, a random walk with transitory shocks provides a fairly accurate description of the dynamics of log wages. Using the values in Table 2.4, we have $\sigma_\eta^2 = 0.0130$ and $\sigma_w^2 = 0.0380 - 2 \times 0.0130 = 0.0120$. The standard deviation of permanent innovations is substantial, $\sigma_w = 11\%$ per year. The variance of the transitory shocks σ_η^2 is consistent with the variance of the measurement error for earnings of US males reported by Bound and Krueger (1991). Hence, transitory shocks can be ignored in the analysis of job relocation.

Finally, we inspect whether the variance of the innovations in wages depends on tenure or experience. Table 2.5 present results for the Koenker (1981) Studentized LM version of the Breusch-Pagan (1979) test for homoskedasticity of u_t for both stayers and movers: the squared residuals from (2.20) are regressed on a constant term and the control variables.

Table 2.5: Heteroskedasticity test for wage changes within- and between-jobs

	Within-jobs	**Between-jobs**
Intercept	0.0458**	0.1120**
(std. err.)	(0.0042)	(0.0191)
Tenure	-0.0005	-0.0013
(std. err.)	(0.0004)	(0.0030)
Experience	0.0006*	0.0017
(std. err)	(0.0003)	(0.0015)
Observations	8802	1243
SER[(1)]	0.1861	0.3743
R^2	0.0006	0.0011
Breusch-Pagan χ^2 test	$\chi^2_{(2)}$: N*R^2=5.28	$\chi^2_{(2)}$: N*R^2=1.37

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

(1) SER= standard error of the regression (root mean square error)

First, in a joint test for tenure and experience, the null hypothesis of homoskedasticity cannot be rejected for both stayers or movers. At the level of single variables, only the effect of experience for job stayers is marginally significant. A learning model would imply a higher variance early on in the job, when the firm still has to learn the capability of the worker, see Jovanovic (1979b) and Topel and Ward (1992). The results reported in Table 2.5 do not support this idea. The variance of wage changes is substantially higher for job movers than for stayers. This result might be explained by the strong fall in wages in the last year before separation, see Figure 2.2 and Proposition 2 above. Alternatively, it might be due to search frictions, so that workers cannot collect the best alternative job option, unlike we assume in this paper.

2.3.4 Wage dynamics

Table 2.6 presents estimation results for equations (2.11) and (2.12), both for the whole sample and separately for the complete and incomplete job spells, and respectively for the job transitions, with heteroskedasticity-robust standard errors[14]. We add as an additional control experience, to account for concavity in the experience profile, which has been documented in Section 3.3. As long as the concavity in the experience profile affects p_t and r_t in the same way, it hardly affects the theoretical structure of our model.[15] The theoretically relevant regressors, $\Delta \mathrm{E}(\Omega|\cdot)$ and Ω, have the right sign in all specifications. The coefficients in the regression for job switchers in column 4 are badly determined. This is not surprising since Ω, the intercept, and the coefficient on experience are highly collinear (since Ω varies by initial experience, see Table 2.2). Hence, the subsequent discussion on Ω focuses on the first of the three columns. The coefficient on $\Delta \mathrm{E}(\Omega|\cdot)$, $(1-\gamma)\,\overline{\sigma}$, varies substantially between the three specifications, and is statistically insignificant in the regression for completed job spells in column 2. Although the differences in the estimated value across columns are insignificant, these results suggest that there is downward rigidity in wages. Where the model predicts a fall in wages at the job compared to the outside wage at the end of a job spell, see Figure 2, the data do not seem to support this idea.

Table 2.6: Wage change regressions: Overall, completed spells, incomplete spells, job transitions

	1: Overall	2: Completed	3: Incomplete	4: Job Switch
Intercept	0.047**	0.049**	0.031**	-8.253
(st. errors)	(0.004)	(0.009)	(0.008)	(6.738)

Continued on next page...

[14]We tested for the absence of individual specific effects in all regressions.

[15]Since the concavity of the experience profile affect p_t and r_t in the same way, it drops out in $b_t = p_t - r_t$. Hence, as long as separations are driven by b_t falling below a constant separation threshold b^T, this concavity does not matter for our analysis. Strictly speaking, the Bellman equations (2.2) do not apply, since the concavity in the experience profile adds another state variable, and hence the separation threshold b^T will depend on t. We ignore this effect.

... table 2.6 continued

ΔE(Ω|.)	0.011*	0.007	0.035*	0.138*
(st. errors)	(0.005)	(0.007)	(0.016)	(0.068)
Ω	0.203**			28.293
(st. errors)	(0.064)			(22.252)
Experience	-0.0016**	-0.0015**	-0.0013**	0.0403
(st. errors)	(0.0002)	(0.0006)	(0.0003)	(0.035)
Observations	8994	1648	6833	513
SER	0.209	0.185	0.197	0.369
R^2	0.008	0.006	0.006	0.024

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

As further check of the notion of downward rigidity, we re-estimated the equation for all completed job spells, eliminating the last within spell wage change before separation.[16] Then, the coefficient for $\Delta E(\Omega|\cdot)$ goes up to 0.037 (std. err. 0.011) in column 1 and 0.043 (std. err. 0.023) in column 2. So indeed, the last observation before a job change, where $\Delta E(\Omega|\cdot)$ is negative to accommodate the downward part of the hump shape profile, does not fit the model well, consistent with the idea of downward rigidity. The variable Ω only shows up in the equation for job switchers, compare equations (2.11) and (2.12). Hence, its coefficient, $\overline{\sigma}$, is only identified by comparing wage changes within job spells and between job spells, that is, in column 1, by identifying the intercept and the coefficients on $\Delta E(\Omega|\cdot)$ and experience on wage changes within job spells, and then identifying the coefficient on Ω as the differential impact on wage changes for job movers.

The estimation results from column 1 imply $\gamma = 1 - 0.011/0.203 = 0.946$. Apparently, separation is driven by selectivity in the shocks to the worker's reservation wage r_t, not to the current job's productivity p_t, which seems somewhat counter-intuitive. Just as a way to illustrate the sensitivity of this estimate to the occurrence of downward rigidity, we redo this calculation for the alternative specification, where we omit the last observation on Δw_t before job change; then $\gamma = 1 - 0.037/0.203 = 0.818$.[17] We can calculate the return to tenure, $\overline{\sigma}\pi = 0.203 \times 0.14 = 2.8\%$ (taking the estimated mean value of $\pi = 0.14$ from Table 2.2 above). However, the high value of γ implies that most of the return to tenure, almost 95%, takes the form of the log reservation wage r_t falling, instead of the inside wage w_t rising, c.f. equation (2.10). The tenure profile due to the rise in log productivity in the current job p_t is really small, $(1-\gamma)\overline{\sigma}\pi = 0.011 \times 0.14 = 0.15\%$

[16] The wage change Δw_T^* is included in the regression for job switchers, so the last change included in the regression for completed spells is Δw_{T-1}. In the alternative specification, these observations are excluded.

[17] This calculation is hazardous, since downward wage rigidity is also affecting the value of $\Delta E(\Omega|\cdot)$ in the regression for job switchers, and hence our estimate of the coefficient for Ω.

(or $0.037 \times 0.14 = 0.52\%$ in the alternative specification). Although the discussion on downward rigidity has shown that this is a rather thin line of identification of γ, this is the first research to actually account for selectivity in the realised outside wages.

Table 2.7 below presents the estimation results for equation (2.14), the variance of σ_z^2. Column 1 reports the coefficients for equation (2.13). Column 2 regresses the squared residuals from the first column on an intercept and the completed tenure. We restrict the sample to jobs lasting more than 1 year, since wage changes for jobs that last less than one year are noisy. The intercept captures the excess variance for job movers, see Table 2.5, and the variance of the transitory shocks in wages σ_η^2, see equation (2.21). The coefficient for tenure is an estimator of $\sigma_z^2 = \Theta^{-1}\text{Var}(w_T^* - w_S)$. However, due to the low sample size, this coefficient is not statistically significant. Using the estimator for σ_w^2 of 0.0120, see Section 3.2, equation (2.14) implies $(1-\gamma)\overline{\sigma} = \sqrt{0.0120 - 0.0104} = 0.040$. Given the noisiness of the estimate of σ_z^2, this test does not have much power. However, the point estimate is consistent with the estimates of $(1-\gamma)\overline{\sigma}$ derived from Table 6, in particular in the alternative specification where we allow for downward rigidity in wages.

Table 2.7: Regression on changes in initial wages between jobs

	1: Initial wages	**2: Residual variance**
Intercept		0.130**
(st. errors)		(0.030)
Tenure	0.033**	0.010
(st. errors)	(0.011)	(0.008)
$\Delta\Omega$	0.168	
(st. errors)	(0.142)	
Observations	328	328
SER	0.400	0.355
R^2	0.121	0.003

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%
We use completed jobs that last more than 1 year (Tenure >1)
The dependent variable in column 2 is the residual variance from column 1

A final question we ask is to what extent the option to switch jobs limits the growth of the variance in log wages over time. Without the option to switch jobs, the variance of log wages would increase linearly over time, due to the fact that z_t and b_t follow a random walk. However, the option to switch jobs allows the worker to eliminate bad trajectories of b_t, thereby compressing its variance. This can be seen from the distribution of incomplete tenures, see Figure 2.6 above, showing that a substantial fraction of the jobs has an incomplete tenure of less than 32 years. There are two mechanisms that lead to compression. First, many jobs have an incomplete tenure of less than 32 years and hence a smaller variance, since the variance increases proportional to incomplete tenure. Second,

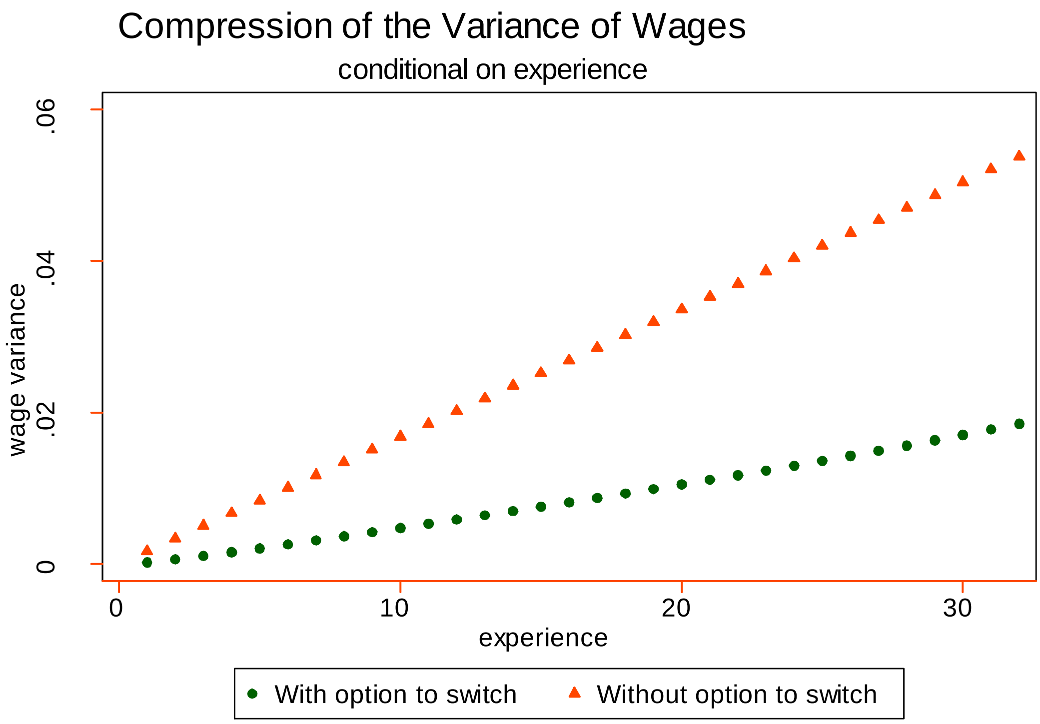

Figure 2.7: Compression effect of the exit option on the variance of b_t

those jobs that are still going on after some period are a selective sample of all the trajectories that have started initially, namely those which never crossed the separation threshold. This selection process compresses the variance. We use the computed density of incomplete tenures from Figure 2.6, and the density of $\Omega_\tau = b_t/\sigma$ conditional on the incomplete tenure τ, $g\left(\Omega_\tau, \tau\right)/\overline{F}\left(\tau\right)$, see equations (2.6) and (2.7). In Figure 2.7, we plot together 2 graphs: first, the evolution of the variance of b_t without the option to switch jobs, the line $\sigma^2\Theta$, and second, the evolution of the variance with that option. The plots reveal that the option to switch jobs compresses the variance of b_t considerably: by almost 65%, after 32 years of experience[18]. Note however that the variance of z_t remains unaffected by this process, while this accounts for the main share of the total variance of w_t in our calculations.

[18]For the computation we use $\sigma = 0.040$, as computed above in the variance comparison test. Hence, after 32 years of experience, without the exit option the variance of b_t would be $32 * 0.040^2 = 0.0512$. With the option to exit the wage variance is compressed after the same period to about 0.018.

2.4 Discussion and conclusions

We analyzed a model of the evolution of wages and the duration of job spells featuring frictionless labor market at the moment of job start –enabling workers to pick the best job alternative at that moment– specific investment and hence subsequent lock-in on the current job, and efficient bargaining over the match surplus. This model explains the data on the job tenure distribution and wages for the USA surprisingly well, the main deviation being that the data suggest there is downward wage rigidity for which we do not allow in our model. We have proven the remarkable result that in this model the evolution of log wages in completed job spells does not provide any information whatsoever on wage-tenure profiles, since this evolution is independent of the drift in log wages. Hence, the tenure profile can only be estimated either from the distribution of tenures or from log wages in incomplete job spells. We have verified that the wage dynamics within jobs closely resembles a random walk; that the predicted job hazard rate is humped shaped with the peak very early in time, closely tracing the empirical evidence on job exits; and that the variance of the within-job wages does not diminish with tenure or experience, a fact that is less easily squared with the learning model. We have further shown that the concavity in the observed tenure profile is easily explained by the selection of the surviving employment matches, even when the underlying tenure profile is linear. In general, the selection effect tends to be much more important than the deterministic trend. This is in fact the first research that looks at selectivity in the realised outside productivities. Remarkably, job separation is driven more by the selectivity in the outside productivity r_t, than by shocks to the inside productivity in the job p_t. Almost 95% of our estimated tenure profile is accounted for by the selectivity in the outside productivity. However, identification of the part of the variance due to variation in r_t is fragile. A rough calculation suggest that allowing for downward rigidity in wages reduces this estimate to 81%. We find excess variance of wage changes at job transition. This might indicate that our assumption of frictionless market for job alternatives at the moment of job change is incorrect.

Appendices Chapter 2: Conditional expectation of Ω_τ

2.A Completed spells: proof of Proposition 2

For the subsequent derivations, it is useful to add the parameter for initial surplus, Ω, as an argument to the survival function of job tenures in equations (2.7) and (2.8), thus $\overline{F}(\tau,\Omega)$ and $f(\tau,\Omega)$. Let $h(\omega,\tau,\Theta)$ be the density of $\Omega_\tau=\omega$ conditional on $A(\Theta)$. Comparing this density to $g(\omega,t)$, there is one additional condition: $\Omega_\Theta=0$. Hence, $h(\omega,t,\Theta)$ can be calculated by applying Bayes's rule. Since Ω_τ is a martingale, the distribution of Θ conditional on $\Omega_\tau=\omega$ is equal to the distribution of $\Theta-\tau$ conditional on $\Omega=\omega$. Hence, its density is $f(\Theta-\tau,\omega)$. Then $h(\omega,\tau,\Theta)$ can be calculated from $f(\cdot)$ and $g(\cdot)$ by Bayes's rule:

$$h(\omega,\tau,\Theta)=\frac{f(\Theta-\tau,\omega)g(\omega,\tau)}{\int_0^\infty f(\Theta-\tau,x)g(x,\tau)\,dx}$$

Substitution of equation (2.6) in the above yields:

$$\begin{aligned}
h(\omega,\tau,\Theta) &= \frac{\omega}{\Omega m(\tau)}\sqrt{\frac{1}{m(\tau)\tau}} \\
&\quad\times\left[\phi\left(\sqrt{\frac{1}{m(\tau)\tau}}\left[\omega-m(\tau)\Omega\right]\right)-\phi\left(\sqrt{\frac{1}{m(\tau)\tau}}\left[\omega+m(\tau)\Omega\right]\right)\right] \\
m(\tau) &\equiv \frac{\Theta-\tau}{\Theta}
\end{aligned}$$

Hence, $\mathrm{E}(\Omega_\tau|A(\Theta))$ satisfies:

$$\begin{aligned}
\mathrm{E}(\Omega_\tau|A(\Theta)) &= \int_0^\infty \omega h(\omega,\tau,\Theta)d\omega \\
&= 2\sqrt{m(\tau)\tau}\phi\left(\sqrt{m(\tau)/\tau}\Omega\right)-\left(\frac{\tau}{\Omega}+m(\tau)\Omega\right)\left[1-2\Phi\left(\sqrt{m(\tau)/\tau}\Omega\right)\right]
\end{aligned}$$

The first and second derivatives of $\mathrm{E}(\Omega_\tau|A(\Theta))$ read:

$$\begin{aligned}\frac{d\mathrm{E}(\Omega_\tau|A(\Theta))}{d\tau} &= -2\sqrt{\frac{1}{m(\tau)\tau}}\phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right)+\left(\frac{1}{\Omega}-\frac{\Omega}{\Theta}\right)\left[2\Phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right)-1\right]\\ \frac{d^2\mathrm{E}(\Omega_\tau|A(\Theta))}{d\tau^2} &= -\sqrt{\frac{1}{m(\tau)\tau}}^3\phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right)\end{aligned}$$

■

2.B Incomplete job spells

Let $h^*(\omega,\tau,\Psi)$ be the density of $\Omega_\tau=\omega$ conditional on $B(\Psi)$. Application of the Bayes rule yields:

$$h^*(\omega,\tau,\Psi)=\frac{\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)}{\int_0^\infty\overline{F}(\Psi-\tau,x)g(x,\tau)\,dx}$$

Hence, $\mathrm{E}(\Omega_\tau|B(\Psi))$ satisfies:

$$\begin{aligned}\mathrm{E}(\Omega_\tau|B(\Psi)) &= \int_0^\infty\omega h^*(\omega,\tau,\Psi)d\omega\\ &= \frac{\int_0^\infty\omega\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}{\int_0^\infty\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}\end{aligned}$$

where $\overline{F}(\Psi-\tau,\omega)$ is given by equation (2.7). This expression is evaluated numerically.

Chapter 3

Returns to Tenure or Seniority?

3.1 Introduction

Why does Lars earn a lower wage than Jens, while they both do exactly the same job, at the same firm, and with equal skills? And why is Pedro fired when his employer has to scale down employment and his colleague Miguel allowed to stay at the firm, while again they do the same job? Some might think that the answer to these questions is obvious: it is simply because Jens and Miguel have a longer tenure at the firm than Lars and respectively, Pedro. Nevertheless, we do not know of any paper within economics that establishes and justifies these regularities. This paper seeks to fill this gap and to provide a simple explanation for the occurrence of these phenomena. Using matched worker-firm data for Denmark and Portugal, we show that a worker who is hired last, is likely to be fired first (Last In, First Out; LIFO). Analogously, we show that there is return to *seniority* in wages. In both cases, our claims are different from saying that your *tenure* at the job affects negatively your job exit hazard or, respectively, affects positively your wage. Seniority is different from tenure in that it measures the worker's tenure *relative* to the tenure of her colleagues. Your seniority is your rank in the tenure hierarchy of the firm. Hence, we need all-encompassing matched worker-firm data to establish a worker's seniority because we need to know the tenure of all the firm's workers. Thus, when we claim that seniority affects your separation risk, we mean that on top of the negative duration dependence of the hazard rate, being a senior worker with many more junior colleagues has a further negative effect. Similarly, when we claim that there is a return to

[0]This chapter is based on Buhai, Portela, Teulings and Van Vuuren (2008). We are grateful for insightful comments and suggestions to Marianne Bertrand, Paul Bingley, Juergen Maurer, Espen Moen, Dale Mortensen, Kevin Murphy, Jean Marc Robin, Robert Topel, Valerie Smeets, Lars Vilhuber and participants in seminars and conferences at LMDG 2007 in Sandbjerg, EALE 2007 in Oslo, ESEM 2007 in Budapest, University of Mannheim, Aarhus School of Business, ESPE 2006 in Verona, SOLE 2006 in Boston and CAFE 2006 in Nuremberg. The usual disclaimers apply.

seniority in wages, we mean that on top of the return to tenure as usually measured, there is return to seniority. We offer a simple economic theory of why firms and workers would agree on applying a LIFO layoff rule and why that leads to a return to seniority in wages. A LIFO layoff rule is a way to protect the interests of incumbent insiders when hiring and training new workers. Without this protection, the incumbents would have an incentive not to train any new worker. The LIFO layoff rule provides protection against layoff for senior workers, and hence gives these workers additional power to bargain for a higher wage, leading to a return to seniority. To the extent that this return to seniority is a compensation for the worker bearing part of the cost of specific investment in the relation between the worker and the firm, the LIFO rule can be interpreted as a protection of the worker's property right on her specific human capital in the relation with the firm. We show that worker turnover is maximal and the expected job duration is minimal when the surplus and the cost of the specific investments are shared between the worker and the firm in the same proportions, which is an application of the Hosios (1990) condition. In some sense, stringent Employment Protection Legislation (EPL) acts as an artificial way to increase the specific investment in the relation, thereby reducing turnover and increasing the expected job duration. Comparing Denmark and Portugal, we see that Portugal has much more stringent EPL than Denmark, and in accordance with our theoretical predictions, a much higher expected job duration than Denmark.

Our theory is based on a dynamic model of the firm with stochastic product demand and irreversible specific investments for each newly hired worker, similar to Bentolila and Bertola (1990). Dixit (1989) considers the same model, but then for an individual worker. Labor demand follows a geometric random walk in these models. Bentolila and Bertola calculate the optimal hiring and firing points, by considering, for the current employment level, the expected discounted marginal revenue of hiring an additional worker, accounting for the expected moment when it is efficient to fire that worker, taking as given all workers currently employed by the firm and disregarding any workers that might be hired in the future. In this way, the hiring and firing of each worker can be considered separately of the hiring and firing of all other workers, transferring a firm level model into a model of an individual worker, as in Dixit (1989). This turns out to be equivalent to applying a LIFO separation rule. Whereas Bentolila and Bertola (1990) and Dixit (1989) take wages as given, we allow for wage bargaining over the surplus generated by the specific investment. Here, we apply an idea developed by Kuhn (1988) and Kuhn and Robert (1989). They start from the distinction in trade union theory between the right-to-manage model, where the union bargains for wages above the market wage and the firm reduces its labor demand in response to this higher wage (it has the right to manage) -leading to an inefficiently low employment- and the efficient bargaining model, where the union and the firm bargain simultaneously over wages and employment, so that employment remains at its efficient level. Kuhn and Robert observe that there is an alternative way for workers to extract rents from the firm, while retaining both the right-to-manage feature and efficiency in employment setting. Their idea is to bargain for a layoff order and for a wage schedule where inframarginal workers get higher wages than marginal workers. The firm cannot fire the expensive inframarginal workers without first firing the cost effective marginal workers. When this wage schedule is properly set, the firm will pick the efficient employment level. As a consequence of this setup, equally productive workers receive different wages, just based on their position in the layoff order, just like Lars and Jens in the opening sentence of this paper. Kuhn and Robert elaborate their ideas in a static framework. Here, we introduce them in the dynamic model of Bentolila and Bertola, leading to a return to seniority in wages. We take an eclectic approach, that is, we do not start from an explicit bargaining game, but from positing a log linear sharing rule of the surplus of the specific investment. However, we impose one feature that characterizes Nash bargaining, namely efficient bargaining: as long as there is a surplus, the worker and the firm will be able to agree on a distribution of that surplus that makes continuation of the relation mutually beneficial. This guarantees that there is efficiency on the firing side. However, the efficiency of hiring decisions depends on a different issue, namely the Hosios condition, which requires the surplus generated by the specific investment to be shared between the worker and the firm in the same proportions as their shares in the cost of the investment. If not, hiring is below the efficient level due to a hold up problem. We elaborate our model under the assumption that the firm must pay for the full cost of the specific investment, so that any return to seniority implies sub-efficient hiring. Under risk neutrality, contractibility of either specific investment or wages suffices to achieve efficiency, since we can always satisfy the Hosios condition by using one to match the other. When workers are risk averse, any return to seniority is inefficient, as it assigns the worker a risky return that can better be assigned to the risk neutral firm. As an extension, we consider the effect of firing cost, accounting for its upward effect on wages.[1] By the efficient bargaining assumption and the Coase theorem, firing cost does not affect firing, but further deteriorates hiring. Finally, we consider the role of trade unions in this model. The ideas in Kuhn (1988) and Kuhn and Robert (1989) seem to suggest that the return to tenure should be higher in unionized firms, since unions are predicted to use the tenure profile as a rent extraction mechanism. This turned out to be counter-factual: unionized

[1] In Bentolila and Bertola's (1990) analysis of the effect of firing costs, wages are fixed. Accounting for the effect on wage setting turns out to be important for the conclusions.

firms generally have a lower return to tenure, not a higher return, see for instance Teulings and Hartog (1998: 225). We observe that this fits our theory. Incumbent workers have sufficient bargaining power to extract returns to their seniority even in the absence of a formal union: their cooperation is indispensable when the firm wants to transfer the tacit knowledge to newly hired workers. The LIFO layoff rule allows for a decentralisation of the bargaining process, leading to higher wages for senior workers. Instead, the political process within a union would lead to a more egalitarian distribution of the rents among the workers, that is, to higher wages but a lower wage return to seniority.

In the empirical part, we establish a number of features of our model. We show that seniority is an important determinant of job separation. Junior workers have a larger separation probability than senior workers. This effect comes on top of the duration dependence of the hazard, that is, in addition to the fact that the separation probability declines with the elapsed tenure at the job. Second, we show that there is a wage return to seniority. Starting from the seminal papers by Altonji and Shakotko (1987) and Topel (1991), there is a large and still flourishing literature on the estimation of the wage return to tenure. The problem in this literature is that within a job spell, tenure is perfectly correlated with experience. Hence, the first order term of this return can only be estimated using variation between job spells, but that introduces all kind of selectivity problems, which this literature sets out to resolve. This problem is absent in the estimation of the return to seniority, since seniority is not perfectly correlated with experience. Seniority increases for example because new workers enter the firm. From that perspective, changes in seniority are correlated with changes in firm size, since an increase in firm size requires new workers to be hired and, hence, the seniority of the incumbents to increase. Luckily however, seniority is not perfectly correlated to firm size, since then the return to seniority could not be disentangled from the firm size wage effect: seniority does also increase by more senior workers leaving the firm, for example due to retirement. In our regressions, we use within job spell variation and we include both tenure and firm size as controls. Nevertheless, we are still able to find wage returns to seniority of 1 to 2 % in Portugal, and returns half that range in Denmark. Including seniority reduces the coefficients for tenure and firm size by 5-30 %, suggesting that tenure and firm size served at least partly as proxies for seniority in previous regressions. The return to seniority turns out to be of the same order of magnitude for males and females, but much larger for high- than for low-educated workers. Our theory also implies that returns to seniority are higher in industries with a high degree of monopoly power. We make an attempt to test this prediction, but our explanatory variables proxying for monopoly power are not strong enough to find an effect.

The paper is set up as follows. Section 2 presents our theoretical framework. In Section 3, we describe the data and the relevant labour market institutions in Denmark and Portugal, and we present our estimation results. Section 4 summarizes and concludes.

3.2 Theoretical framework

3.2.1 Setup

The model of Bentolila and Bertola (1990) provides a nice starting point for our analysis. Firms face a stochastic iso-elastic demand curve for their output, in logs:

$$n_t = z_t - \eta p_t, \tag{3.1}$$

where $\eta > 1$ is the price elasticity of demand, n_t is log demand, p_t is its log price, and z_t is a market index capturing the exogenous evolution of demand; z_t is assumed to follow a Brownian with drift, such that $\Delta z \sim N(\mu, \sigma^2)$. Labor is the only factor of production. The production function exhibits constant returns to scale. Without loss of generality, productivity is normalized to unity, so that output is equal to employment. In the model of Bentolila and Bertola (1990), hiring and firing of workers is costly. At this stage we focus on hiring cost, denoted by I. This cost is interpreted more broadly as the specific investment that has to be made by the firm at the start of an employment relationship. It is irreversible: once made, the cost cannot be recouped by ending the employment relation. For simplicity, we assume that this investment can be made instantaneously, so that no time elapses between the start and the end of the investment process. At the outside market, workers can earn a reservation wage, which is constant over time. It is most convenient to think of this reservation wage as the return to self employment. Without loss of generality, it is normalized to unity. Hence $w^r = 0$, where w^r denotes the log reservation wage. We assume both workers and firms to be risk neutral.

As a benchmark, we analyze first the simple case where firms pay workers their reservation wage and where there are no specific investments required for starting an employment relationship, $I = 0$. In that case, labor demand can be adjusted costlessly at each point in time. Hence, the optimal strategy is to maximize instantaneous profits Π_t:

$$\Pi_t = e^{n_t}\left(e^{p_t} - 1\right),$$

subject to the demand curve (3.1). The first order condition implies

$$\begin{aligned} p_t &= \pi, \\ mr\left(z_t - n_t\right) &= \frac{1}{\eta}\left(z_t - n_t\right) - \pi, \\ n_t &= z_t - \eta\pi, \\ \pi &\equiv \ln \frac{\eta}{\eta - 1} > 0. \end{aligned} \tag{3.2}$$

where $mr\left(z_t - n_t\right)$ denotes the log of the marginal revenue for the firm of hiring an additional worker, conditional on the state of demand z_t and log employment n_t. The parameter π is the log of the ratio of price over wage cost, when marginal cost and marginal revenue are equal. This ratio is greater than unity due to the monopoly power of the firm at the product market. The firm's price is constant over time, while its labor demand follows a random walk. The latter implication is consistent with Gibrat's law that tends to hold for large firms, see for instance Jovanovic (1982).

Next, consider the optimal strategy with specific investments, $I > 0$. Then, labor demand cannot be adjusted costlessly. On the hiring side, an additional worker requires a specific investment, which has to be recouped from future profits. Moreover, this investment is irreversible, so that delaying hiring has an option value. On the firing side, firing per se is costless, but irreversible. If demand surges after having fired the worker, the firm is unable to benefit from that demand without incurring the cost of the specific investment again. Hence, retaining the worker has an option value, too. Bentolila and Bertola (1990) show that the optimal policy of a firm is to hire workers whenever p_t reaches a constant upper bound $p^+ > \pi$ and to fire them whenever p_t reaches a lower bound $p^- < \pi$. The hiring bound p^+ exceeds π due to the necessity of the firm to recoup the cost of specific investments and due to the option value of postponing hiring, while the firing bound p^- is below π due to the option value of postponing firing. The situation is sketched in Figure 1. below.

The present employment level is denoted by n_0 and the present market index by z_0. If the market index rises above z^+, the firm hires additional workers to avoid p rising above p^+. If the market index falls below z^-, the firm fires workers to avoid p falling below p^-. Hence, p_t follows a random walk between p^- and p^+, while n_t is constant in this interval. However, when p_t drifts outside these boundaries, the firm uses n_t as an instrument to control p_t. Then, p_t is held constant, and n_t starts drifting, either up (if $p = p^+$), or down (if $p = p^-$). Bentolila and Bertola (1990) provide expressions for both boundaries.

Suppose we impose a LIFO separation rule upon this firm. We can index each worker by the log employment level of the firm at the date the worker is hired. A worker hired

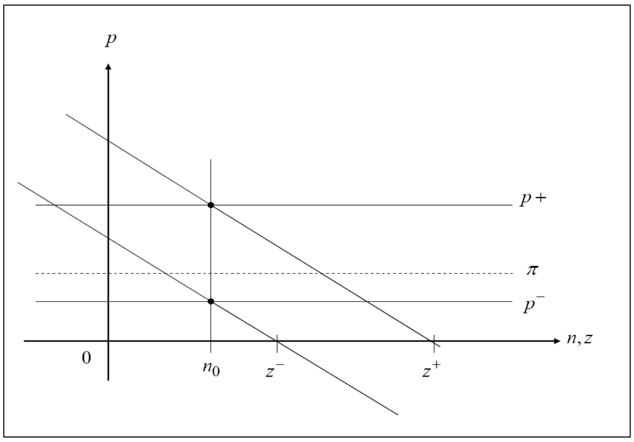

Figure 3.1: Firing-hiring boundaries with stochastic market index

at time h gets rank q, $q = n_h = z_h - \eta p^+$. Her seniority index at time t is defined as $n_t - q$. The less senior the worker, the shorter her tenure, and the lower her seniority index $n_t - q$. Hence, the most senior worker $q = 0$ has seniority index $n_t - q = n_t$, while the least senior worker at time t with $q = n_t$ has seniority index $n_t - q = 0$. The LIFO layoff rule implies that a worker hired at time h with index $n_t - n_h$, will be fired at the first moment $f > h$ that employment is back at the level n_h and $p_f = p^-$, that is, when $z_f - z_h = -\eta\,(p^+ - p^-)$. By construction, f is the first point in time that z_t has travelled down a distance $\eta\,(p^+ - p^-)$ from its initial value z_h. Whether or not new workers have been hired after time h by z_t rising above z_h for some $t, h < t < f$, and if so, how many, is immaterial to this conclusion, since these workers are indexed $n_f - n_t < n_f - n_h$, and hence, the LIFO separation rule imposes that these workers will be fired before the worker with seniority $n_f - n_h$. Bentolila and Bertola's model of firm level employment supplemented with a LIFO layoff rule corresponds one-to-one with a simple model of individual job tenures. In this model, a worker hired at time h, with $z_t = z_h$, will be fired a the first time that z_t has travelled down a distance $\eta\,(p^+ - p^-)$; see Buhai and Teulings (2006) for a recent elaboration of that model.

3.2.2 Rationale for LIFO

Why would a firm use a LIFO layoff strategy? In the simple world discussed above, where the firm pays the worker her reservation wage, there is no rationale for such a rule. Since the worker receives her reservation wage, she is indifferent between working at the firm or being laid off. Hence, there is no point in fixing an order of layoff. However, if we relax the assumption that the firm pays its workers their reservation wage and we attribute incumbent workers some bargaining power, the quasi rents of the specific investment enable these workers to capture wages above the reservation wage. In that case, a layoff order carries practical relevance, as it protects the 'rights' of senior workers (those who are hired first). Kuhn (1988) and Kuhn and Robert (1989) offer a neat further legitimation for using such a rule. Their idea is based on the classic distinction in the theory of unionized wage setting between right-to-manage or labour demand curve models, on the one hand, and efficient bargaining models, on the other hand. In the former, unions raise wages above the reservation wage. However, the firm can set employment unilaterally ("the firm has the right to manage"). From the point of view of the firm, wages are equal to marginal cost. Hence, the higher wage rate reduces the firm's labour demand. This outcome has been criticized for leaving gains from trade between the firm and the workers unexploited. Additional workers would be willing to join the firm at a wage rate between the reservation wage and the negotiated wage, and the firm would be willing to hire them at that wage. By bargaining on wages and employment simultaneously, the firm and the union can exploit these gains from trade. Kuhn and Robert's idea is that these gains from trade can also be exploited, while maintaining the right-to-manage feature that wages are negotiated between the firms and the union (or: its workers) and that the firm sets employment unilaterally. This can be done by fixing the order of layoff, and differentiating wages by the position in the layoff order. A firm can only fire senior workers after having first fired all junior workers. Senior workers earn the highest wage since they can only be fired after all juniors been fired and therefore they feel sufficiently protected to demand higher pay, any resultant job loss falling on their less senior colleagues. This is a form of price discrimination on the side of the union. First degree price discrimination results when the union has full bargaining power. Inframarginal senior workers receive part of their inframarginal productivity surplus.

Kuhn and Robert (1989) specify their theory in static framework. In that case, the layoff ordering can be based on any variable, height, IQ, experience, or what else springs to mind. Combining the model of Bentolila and Bertola (1990) with a LIFO layoff ordering provides a straightforward way to cast the ideas of Kuhn and Robert in a dynamic framework. Then, the prevalence of a LIFO ordering has a natural economic interpretation. The senior workers' future wage claims are sensitive to the firm hiring new workers, since after the specific investments have been made, these new workers are perfect substitutes for senior workers. The firm could in principle hire new workers for a low wage, and fire the senior workers instead. The lack of commitment on the side of the firm of not benefitting from this strategic option has an adverse effect on the set of feasible contracting arrangements open to the firm and its workers. Suppose that the specific investment of new workers is largely made up from acquiring the tacit knowledge of the firm's production process and the transfer of this knowledge can be blocked by senior workers, or suppose that senior workers can harass newcomers, as suggested by Lindbeck and Snower (1990). In that case, hiring new workers requires the consent of senior workers. At the same time, the firm has a commitment problem: how can it credibly promise senior workers not to use new workers as a replacement for them, after the transfer of the tacit knowledge to the new workers is completed? Due to this commitment problem, gains from trade from hiring new workers cannot be exploited. A LIFO separation rule is a solution to this commitment problem, by providing senior workers protection against being laid off before newly hired workers, so that there are no disincentives to cooperate in training new workers.

3.2.3 Wage sharing rule

We operationalize the LIFO idea by positing a linear relation for the log wage as a function of the current state of product demand z_t and the seniority index $n_t - q$:

$$w(z_t - q) = \omega + \frac{\beta}{\eta}\left(z_t - q - \eta p^-\right) \tag{3.3}$$

where $0 < \beta < 1$ and where ω is the log wage at which a worker is indifferent between remaining employed at the firm and being laid off, given the LIFO layoff rule and equation (3.3). The parameter ω can therefore be interpreted as the reservation wage of an incumbent worker. We shall derive an expression for it below, when discussing the worker's problem. At the moment of firing the worker with seniority index $n_t - q$, it must be true that $z_t = q + \eta p^-$, by the definition of the firing bound. The factor $\eta^{-1}\left(z_t - \eta p^- - q\right)$ is equal to $mr\left(z_t - q\right)$, the log of surplus of marginal revenue above marginal cost, compare expression (3.2), conditional on employment being equal to the worker's seniority index $n_t - q$; that is, if there were no workers in the firm with higher seniority, or equivalently, if there were nobody in the firm with a lower tenure than worker $z_t - q$. This marginal revenue is a counterfactual, in the sense that actual employment can be larger, $n_t \geq q$. We can therefore just as well refer to this "marginal" revenue as the "infra marginal" revenue, because it would be the marginal revenue only if the firm were first to fire all workers with a higher q. Equation (3.3) implies that senior workers receive a share β of this surplus of log inframarginal revenue above the log marginal outside option. The parameter β can be interpreted as the bargaining power of workers, though strictly speaking this interpretation lacks a foundation in a formal bargaining model.[2] The log linearity of equation (3.3) is just imposed for the sake of analytical convenience. Equation (3.3) implies efficient bargaining: as long as there is a positive surplus both parties get a share of it, so that it is rational for both to continue the employment relation. As soon as the surplus is vanished, separation will occur, which by then is the efficient outcome. Note that equation (3.3) depends on $z_t - q$, not on z_t and q separately. Since z_t is closely related to log firm size n_t (apart from the effect of insulating n_t from fluctuations in z_t whenever p_t is in between the hiring and firing bounds, p^+ and p^-), $z_t - q$ can be interpreted as an index of seniority relative to firm size. The return to the seniority index (or simpler, to seniority), β/η, is increasing in the bargaining power of the workers, β, and in the monopoly power of the firm on its product market, η^{-1}. Since $0 < \beta < 1$ and $\eta > 1$, the elasticity of wages with respect to the index $n_t - q, \beta/\eta$, must be between zero and unity.

3.2.4 The worker's problem

The value of ω can be derived using the theory of option values, see Dixit and Pindyck (1994). Let $V(z_t - q)$ be the asset value of holding a job at a firm. By Ito's lemma $V(z_t - q)$ satisfies the Bellman equation

$$\rho V(z_t - q) = e^{w(z_t - q)} + \mu V'(z_t - q) + \frac{1}{2}\sigma^2 V''(z_t - q),$$

where ρ is the interest rate, such that $\rho > \mu + \frac{1}{2}\sigma^2$; $e^{w(z_t-q)}$ is the current wage. The relevant solution to this second order differential equation reads

$$\begin{aligned}
V(z_t - q) &= \frac{1}{r(\beta/\eta)} \exp\left[\omega + \frac{\beta}{\eta}\left(z_t - q - \eta p^-\right)\right] + A \exp\left[\lambda^-(z_t - q)\right], \\
r(x) &\equiv \rho - \mu x - \frac{1}{2}(\sigma x)^2 \Rightarrow r(0) = \rho, r(1) = \rho - \mu - \frac{1}{2}\sigma^2 > 0 \\
\lambda^{+,-} &\equiv -\mu/\sigma^2 \pm \sqrt{\mu^2/\sigma^4 + 2\rho/\sigma^2} \Rightarrow \lambda^- < 0, \lambda^+ > 1.
\end{aligned}$$

where we substitute $w(z_t - q)$ for equation(3.3) and where A is a constant of integration that remains to be determined. $r(x)$ is a modified discount rate, accounting for the drift and the variability of z_t; note that $r(0) = \rho$; $r(1)$ must be positive for a bounded solution for the welfare of the worker to exist. Hence, the first term is the net discounted value of expected wage payments, disregarding the worker's option to quit the firm when wages

[2] In the case of a single worker firm, where we could apply the theory of two player bargaining, as in Buhai and Teulings (2006), the log linear sharing rule would be almost equivalent to Nash bargaining, which would yield a linear sharing rule.

fall too far below the reservation wage. The second term, $A \exp\left[\lambda^{+,-} z_t\right]$, is the option value of separation. Only one of the roots $\lambda^{+,-}$ is relevant here. For large values of z_t, the firm is doing well and, hence, keeping the job is attractive for the foreseeable future. The option value of separation must converge to zero, which is the case for the negative root λ^-, since $\lim_{z\to\infty} A \exp\left[\lambda^- z\right] = 0$. Hence, the constant of integration for the positive root λ^+ is equal to zero. Efficient bargaining implies that it is optimal for the worker to separate when $z_t = q + \eta p^-$. Two conditions must hold for that value of z_t to be optimal: the value matching and the smooth pasting condition. The value matching condition states that the asset value of holding the job should be equal to the asset value after separation, that is, the net discounted value of the reservation wage, ρ^{-1}. The smooth pasting condition states that for small variations in z_t, the worker remains indifferent between holding the job and separation, since the worker should not regret separation after a small perturbation of z_t because separation is irreversible. This requires the first derivative of $V(z_t - q)$ with respect to z_t to be zero. Using $z_t - q = \eta p^-$ at the moment of separation, both conditions read

$$
\begin{aligned}
V\left(\eta p^-\right) &= \frac{1}{r\left(\beta/\eta\right)} e^{\omega} + A e^{\lambda^- z} = \frac{1}{\rho}, \\
V'\left(\eta p^-\right) &= \frac{\beta}{\eta r\left(\beta/\eta\right)} e^{\omega} + \lambda^- A e^{\lambda^- z} = 0, \\
\omega &= \ln r\left(\beta/\eta\right) - \ln \rho - \ln\left(1 - \frac{\beta}{\eta\lambda^-}\right) < 0,
\end{aligned} \tag{3.4}
$$

where the final equation follows from the elimination of A in the first two. ω is below the log reservation wage $w^r = 0$ since separation is an irreversible decision. If the demand for the firm's product, z_t, goes up after the separation decision, the worker is no longer able to benefit from the wage increase. Hence, workers are prepared to incur some loss before they decide to separate. The higher the worker's share β in the log surplus, the lower is ω, since expected future revenues are higher so that workers are prepared to accept greater losses before separation. Similarly, ω is declining in the drift μ, since a higher drift raises expected future revenues, and it is declining in the variability of demand σ^2, since a higher variability raises the option value of hoping for a future increase in the surplus.

3.2.5 The firm's problem

Given this wage setting rule, we derive the firm's optimal strategy. We use a methodology that is similar to Bentolila and Bertola (1990). We attribute to each worker her marginal revenue and her wage, taken the employment of workers hired previously as given, and then consider when it is optimal for the firm to hire and subsequently fire this worker. In this way, we can consider the decision to hire and fire the N_t-th worker ($N_t \equiv \exp n_t$) separately of the hiring and firing of workers hired before this worker, and of workers hired afterwards. The model is thus a straightforward extension of Dixit and Pindyck (1994: 216), the only difference being that wages are constant in Dixit and Pindyck, while they vary with the state of demand z_t in this model. Let $F(n_t, z_t)$ be the asset value of the firm for the N_t-th worker. The Bellman equation for $F(n_t, z_t)$ satisfies

$$\rho F(n_t, z_t) = \exp[mr(z_t - n_t)] - \exp[w(z_t - n_t)] + \mu F_z(n_t, z_t) + \frac{1}{2}\sigma^2 F_{zz}(n_t, z_t). \quad (3.5)$$

The first term is the marginal revenue of the N_t-th worker, see equation (3.2), the second term is the wage for that worker. The relevant solution to this differential equation reads

$$\begin{aligned} F(n_t, z_t) &= \frac{1}{r(\eta^{-1})}\exp\left[\frac{1}{\eta}(z_t - n_t) - \pi\right] - \frac{1}{r(\beta/\eta)}\exp[w(z_t - n_t)] \\ &\quad + B^- \exp\left[\lambda^-(z_t - n_t)\right]. \end{aligned}$$

The final term is the option value of separation, with B^- being the constant of integration. As in the case of the worker, the positive root λ^+ is irrelevant, since the option value must converge to zero for large values of z_t (since then the option to fire the worker has no value). Suppose the firm employs less than N_t workers. Then, the option value of hiring the N_t-th worker at some future date reads

$$G(n_t, z_t) = B^+ \exp\left[\lambda^+(z_t - n_t)\right],$$

where B^+ is the constant of integration. There are no current costs or revenues, hence only the option value term matters. Since this option value converges to zero for low values of z_t, here only the positive root λ^+ applies. The value matching and smooth pasting conditions read

$$\begin{aligned} F\left(z_t - \eta p^-, z_t\right) &= G\left(z_t - \eta p^-, z_t\right), \quad (3.6) \\ F_z\left(z_t - \eta p^-, z_t\right) &= G_z\left(z_t - \eta p^-, z_t\right), \\ F\left(z_t - \eta p^+, z_t\right) &= G\left(z_t - \eta p^+, z_t\right) + I, \\ F_z\left(z_t - \eta p^+, z_t\right) &= G_z\left(z_t - \eta p^+, z_t\right). \end{aligned}$$

The first pair refers to the firing decision, the second to the hiring decision. The first condition states that at the moment of firing, when by definition $n_t = z_t - \eta p^-$, the asset value of keeping the worker is equal to the option value of a vacancy. The second equation is the smooth pasting condition, which states that this condition also applies for slight variations of z_t, so that the firm wouldn't regret a decision to fire after a slight variation in z_t. The third equation is the value matching condition for the moment of hiring, when $n_t = z_t - \eta p^+$: the asset value of hiring the worker should be equal to the cost of investment and option value of filling the vacancy at a later point in time. The final equation is the smooth pasting conditions for the moment of hiring. This system of four equations determines four unknowns, the constants of integration, B^- and B^+, and the hiring and firing boundaries, p^- and p^+. The system is non-linear, and its analysis is relegated to Appendix A, where we prove the subsequent results. A unique, economically meaningful solution to this system exists, where $B^- > 0$ and $B^+ > 0$, and where $p^+ > 0 > p^-$. The elimination of B^- from the first two equations of (3.6) yields

$$p^- = \ln r\left(\eta^{-1}\right) - \ln\rho + \pi - \ln\left(\frac{\eta\lambda^- - 1}{\eta\lambda^-}\right) - \ln\left(1 - \rho\frac{\lambda^+ - \lambda^-}{\lambda^-}B^+\exp\left[\eta\lambda^+ p^-\right]\right). \quad (3.7)$$

The firing bound p^- does not depend on β, except for the effect of β on B^+, which is the option value of refilling the vacancy at a later moment. This is an application of the Coase theorem: under the efficient bargaining, the distribution of the surplus of the employment relation does not matter for the actual level of employment. The option value of rehiring comes in because when the firm decides to fire the N_t-th worker, it always has the option to rehire at a later moment. The larger the distance between the hiring and firing threshold, $p^+ - p^-$, the longer it will take (in expectation) before the firm will find it attractive to refill the vacancy, and hence the smaller is the option value associated with that. Keeping constant all other parameters of the model, an increase in the bargaining power of the workers β raises the distance $p^+ - p^-$

$$\frac{d\left[p^+ - p^-\right]}{d\beta} > 0.$$

The higher the workers' bargaining power β, the less volatile will be the employment, since employment is insulated from shocks in demand z_t over a larger interval of $\eta p^- < n_t - z_t < \eta p^+$, and the larger is therefore the expected tenure of a newly hired worker. This is the consequence of a hold up problem. The larger the workers' bargaining power, the higher the hiring threshold p^+, since the firm reaps a smaller share of future surpluses created by the specific investments in new workers, so that a larger initial surplus of marginal revenue above the reservation wage is required to recoup the cost of these investments. Since $p^+ - p^-$ is larger, the option value of rehiring is lower, and hence the firing threshold

p^- is lower, though this indirect effect of β on p^- is smaller than the direct effect on p^+. Hence, a higher bargaining power of workers reduces the firing threshold and postpones separation. This implication squares well with the findings in Bertrand and Mullainathan (2003), who show that when firms are insulated from takeovers, the wages of the incumbent employees are higher, suggesting a higher value of β. This goes hand in hand with lower rates of creation of new plants, which in the context of our model is similar to hiring new workers. Bertrand and Mullainathan also report a lower rate of destruction of old plants, or in the context of our model, a lower firing bound, p^-. The larger is β, the lower is the option value of future rehiring, and hence the less attractive it is to fire a worker.

3.2.6 Explanation of the firm size wage effect?

The firm size effect on wages has been extensively documented, see Brown and Medoff (1989). Can our model offer an explanation for the firm size wage effect? When we look at the issue from the point of view of an individual worker, the evolution of her seniority $n_t - q$ is driven by the evolution of log firm size n_t. In reality workers retire at some point in time. When more senior workers retire, a worker's seniority goes up even at constant firm size. Here, we abstract from retirement, such that firm size is the only driver of changes in seniority. At first sight, this suggests that our theory could explain the firm size wage effect. Nevertheless, this turns out not to be true. The average log wage in a firm at the firing bound p^- satisfies

$$e^{-n_t}\int_{-\infty}^{n_t} w(z_t - q)e^q dq = e^{-n_t}\int_{-\infty}^{n_t} \left[\omega + \beta/\eta\left(z_t - q + 1 - \eta p^-\right)\right] e^q dq = \omega + \beta/\eta,$$

where in the final equality we use the fact that at the firing bound, $n_t = z_t - \eta p^-$. Hence, the average log wage does not depend on firm size. The intuition is that the positive effect of the wage increase for the incumbent workforce is exactly offset by the negative effect of the below average log wage for the new hires. Thus, although this model predicts firm size to be a driver for the changes in the wages of incumbent workforce, it does not explain why wages for the firm as a whole depend on firm size. However, the average log wage does depend on the parameter β/η. Other things equal (in particular, human capital of the workforce), the model predicts the return to seniority, β/η, to be increasing in the average log wage.

3.2.7 Who gets hired and the welfare cost of hold up

To close the model, we have to explain who gets hired by a firm and who does not. The log wage of a worker who is just hired is higher than the wage of a worker who is at the borderline of being laid off, that is

$$w\left(z_t - \eta p^+, z_t\right) > w\left(z_t - \eta p^-, z_t\right) = \omega.$$

Since the asset value of a worker who is on the borderline of being laid off is equal to the net present value of her reservation wage, $1/\rho$, the asset value of a worker who is just hired must be higher than $1/\rho$. Hence, new jobs at the firm are rationed. A convenient way to model this rationing process is to introduce unemployment. A worker who is just laid off has two options. Either she can decide to collect her outside wage by becoming self employed, or she can decide to queue for a new job at a firm. During this waiting period she cannot produce as a self employed. For simplicity, we assume that leisure has no value.[3] New jobs at firms arrive at a rate λ per unit of the labor force and are distributed randomly among the unemployed. Hence, the asset value of unemployment, V^U, satisfies

$$\rho V^U = \frac{\lambda}{u}\left[V\left(\eta p^+\right) - V^U\right],$$

where u is the unemployment rate. λ/u is the arrival rate of a new job for unemployed. The lower unemployment, the higher this arrival rate, since there are less people among whom new jobs have to be distributed. $V\left(\eta p^+\right) - V^U$ is the asset gain of getting a job offer. The level of unemployment follows from the no-arbitrage condition between self employment and unemployment

$$u = \lambda\left[V\left(\eta p^+\right) - \frac{1}{\rho}\right], \tag{3.8}$$

where we use $V^U = 1/\rho$, the asset value of self employment[4]. The higher the asset gain of getting a job at a firm, the higher must be unemployment. Though the efficient bargaining assumption generates efficiency in the layoff decision of firms, it does not achieve efficiency on the hiring side. There are two types of inefficiency.[5] First, not all gains from trade

[3] Allowing for a value of leisure would not change the predictions of model. It would make unemployment less costly per unit of time, but this effect would be exactly offset by the rise in unemployment.

[4] We assume: $u < 1$. If $u > 1$, the outside option of self employment would become irrelevant, and the reservation wage ω and job arrival rate λ would become endogenous. ω would rise till so many workers are fired, and so few workers are hired till λ is such that the no arbitrage condition holds for $u = 1$.

[5] Throughout the paper, we do not pay attention to a third type, the inefficiency caused by the monopoly power of the firm vis-a-vis consumers.

between the worker and the firm are exploited. Firms would hire more workers if $\beta = 0$, since p^+ is an increasing function of β. Firms' perception of the marginal cost of hiring a worker in net present value terms exceeds the social cost by the same amount as the asset gain for an unemployed of getting a job offer, $V(\eta p^+) - 1/\rho$. This gives rise to a Harberger triangle. Next to this Harberger triangle, there are the costs of rationing that dissipate workers' surplus. The no-arbitrage condition (3.8) implies that the workers as a group spoil their whole share in the quasi rents in wasteful unemployment.

The inefficiency problem arises by a violation of the Hosios (1990) condition. Workers are able to capture a share of the quasi rents of the specific investments, whereas they do not bear a corresponding share of the burden of the investment cost. Hence, firms restrain new hiring to push up the net present value of all rents of a new hire above the cost I till their share in the rent suffices to recoup the cost. The inefficiency is due to the non-verifiability of specific investment and the inability of workers to commit on not using their bargaining power after the specific investment has been made. If wages were contractible, workers could commit on not demanding any return to seniority, so that the firm bears the full cost and gets the full revenues of the specific investment, thereby satisfying the Hosios condition. Alternatively, if specific investments were verifiable, the inefficiency would be resolved by shifting some share of the burden of investment to the worker, such that workers bear an equal share of cost of the specific investment as they get from its revenues, again satisfying the Hosios condition. For the latter case, note that the asset value of a worker at the moment of hiring, $V(\eta p^+)$, is independent of the moment of hiring. Hence, although at a particular point in time senior workers get more quasi rents than juniors, at the moment of hiring each worker has the same net present value of expected rents, independent of her rank q, that is, independent of the level of employment at the moment she is hired. Seniors getting higher rents than juniors at a particular point in time reflects the fact that they are able to realize the upside of the risky returns on their share in the specific investment I. Hence, the LIFO separation rule can be interpreted as a protection of their property right on their share in the quasi rents, against the temptation of the firm to fire the expensive senior workers, thereby depriving them from the upside of their risky returns. A LIFO separation rule is then a device for implementing an efficient contract. This argument implies that as long as we do not know what share of the cost of specific investment is born by workers, empirical evidence supporting the relevance of equation (3.3) for wage setting is inconclusive on the issue of whether or not employment is below its inefficient level.

Finally, note that when workers have to pay the full cost of the specific investment, the hold up problem is precisely the reverse. Then, the non-verifiability of workers investment and the inability of firms to commit on not using their bargaining power $1-\beta$ leads to inefficiency. Workers are only willing to enter the firm when the net present value of quasi rents of their investment is so high that their share in this present value suffices to cover the cost of investment. There is a simple statistic enabling the observer to establish which side is overcompensated in the ex post bargaining over the surplus of the specific investment: when workers queue for jobs, so that there is unemployment, firms are held up, as in the basic model; when firms chase after workers, so that there are vacancies, workers are held up.

3.2.8 Extensions

Some extensions to this model are worth discussing. First, relaxing the assumption of risk neutrality on the side of the worker introduces a trade off. As discussed in the previous section, verifiability of the specific investment I is sufficient to implement first best in the standard case with risk neutral workers. With risk averse workers, this conclusion no longer applies. First best requires that workers get paid their reservation wage all the time, and hence that firms bear the full cost of investment. The inability of workers to commit on not using their bargaining power makes first best unattainable in that case. The case of risk averse workers and risk neutral firms is particularly relevant because one can expect capital markets to be much more complete for firms than for workers. It is much easier for the firm to diversify firm specific risks on the capital market than for its workers.

A second extension is the introduction of firing cost, imposed either by law or by trade unions. We think of a firing cost as a wealth transfer W from the firm to the worker at the moment of firing. By the assumption of efficient wage bargaining, this wealth transfer has an impact on the wage bargaining process. The value matching condition of the worker for the moment of firing reads, compare equation (3.4)

$$V\left(\eta p^{-}\right)=\frac{1}{\rho}+W \Rightarrow \omega=\ln r\left(\beta/\eta\right)-\ln\rho-\ln\left(1-\frac{\beta}{\eta\lambda^{-}}\right)+\ln\left(1+\rho W\right)<0. \quad (3.9)$$

The firing cost raises the value of the outside option of the worker by the wealth transfer W. This raises ω. Hence, there are two counteracting effects on the firing bound p^{+}: the direct effect of firing cost makes layoffs less attractive to the firm, while the indirect effect via higher wages makes layoffs more attractive, since workers are more costly due to the higher level of ω. The first order condition for optimal firing now reads, compare equation (3.6)

$$F_z\left(z_t-\eta p^{-},z_t\right)=G\left(z_t-\eta p^{-},z_t\right)+W,$$

where we use the value of ω from equation (3.9) to account for the effect of firing cost on wages. Some calculation, see Appendix A, shows that the value for p^- remains the same as in equation (3.7). The direct and the indirect effect cancel therefore exactly, except for the indirect effect via B^+, the option value of rehiring. Again, this is an implication of the efficient bargaining assumption and the Coase theorem. On the hiring side, firing cost has two effects with the same sign: first, it raises wages via its effect on ω and, second, there is the prospect of having to pay firing cost in case of future layoff. To the extent that workers have excessive bargaining power, in the sense that the net present value of their share in the surplus exceeds the share in the cost of specific investment, which shows up as workers queueing for a job and hence as unemployment, this increase comes to the detriment of efficiency. The paradox here is that firing costs aggravate the unemployment problem that they were meant to resolve. Since the hiring threshold p^+ goes up due to the introduction of firing cost, the expected value of future rehiring is lower, so the firing threshold is lower. Hence, firing cost raises the distance $p^+ - p^-$ and therefore the expected tenure of newly hired workers.

Now that we have discussed these extensions, risk aversion and firing cost, it makes sense to consider the nature of workers' bargaining power. Most economists associate this power with the trade unions.[6] Only unions provide bargaining power to workers. Without unions, firms are supposed to have complete bargaining power. A notable exception is analyzed by Lindbeck and Snower (1990), who point out that the insiders' ability to harass new hires gives them bargaining power vis-a-vis their employer. Without the insiders' consent, firms are effectively unable to introduce new hires. The interesting aspect of Kuhn (1988) and Kuhn and Robert (1989) is that their rank related compensation scheme allows a decentralization of the bargaining process. As soon as the layoff order has been set, each worker can negotiate for herself. When a marginal worker negotiates a wage increase raising her wage above marginal cost, she endangers her own employment, not that of the inframarginal workers.[7] Hence, a LIFO scheme enables workers to exploit their individual bargaining power without workers having to solve their collective action problem. When workers are united in a trade union, more elaborate strategies are available, that yield a higher expected payoff, in particular when workers are risk averse. By trading a higher transfer W, in case of layoff, in exchange for a lower seniority premium, such that the firm's asset value $F(n_t, z_t)$ remains the same, the expected utility for new hires can be improved by shifting the firm specific risk z_t to the firm. The transfer W allows this

[6] This problem has been suggested to us by Kevin Murphy.

[7] The reverse is not necessarily true. An inframarginal worker can bargain a wage above her productivity, if workers with lower seniority capture less than their full productivity. In that case, the firm has an incentive not to fire the inframarginal worker because it first has to fire the marginal worker.

insurance to be extended even beyond the time spell covered by the employment relation with the firm, compare the discussion on the efficient bargaining model of the union. Moreover, the political decision making process within the union, where senior and junior workers have to compromise on the distribution of the rents, is likely to generate support for an egalitarian outcome, as implied by the median voter model.[8] All these arguments suggest that the LIFO model is probably a more appropriate description of a non-union than of a unionized environment. These arguments can also explain, why tenure profiles seem to be flatter in unionized firms, in contrast to what Kuhn (1988) and Kuhn and Robert (1989) seem to predict, see e.g. Teulings and Hartog (1998: 225).

3.3 Empirical framework

The model discussed in the previous section has three testable implications:

1. Gibrat's law: log firm size n_t follows a random walk. This should hold better when the distance between the hiring and firing boundaries is small. We use various standard procedures to test for Gibrat's law for log firm sizes.
2. The Last-in-First-Out separation rule: the workers hired last, leave the firm first. Note that we apply an efficient bargaining model. Hence, the observational distinction between quits and layoffs is arbitrary, compare McLaughlin (1991). As long as there is a positive surplus of the worker's marginal revenue to the firm above the worker's reservation wage, the worker and the firm will strike a deal. As soon as this surplus has vanished, it is in their mutual interest to separate. Whether the separation is initiated by the worker or by the firm is irrelevant. Hence, the model predicts the LIFO separation rule to apply to separations as a whole, not just to layoffs separately. We use duration analysis to test for this implication. Furthermore, we test one cross country implication. Our theoretical model predicts that EPL raises the distance between the hiring and the firing threshold, and hence the expected completed job tenure. Since Portugal has a much stricter EPL, one would expect tenure to be higher there than in Denmark.
3. The dependency of wages on seniority: wages depend on a worker's seniority in the firm, relative to her colleagues, see equation (3.3). We use wage regressions, both in levels and in first differences, to test this implication. Moreover, the higher the monopoly power of the firm, η^{-1}, the higher should be the return to seniority. We use variation in the estimated return to seniority between industries to test this

[8] Equation (3.3) implies a Pareto distribution of wages within the firm, which is heavily skewed to the right.

implication.

The challenging aspect of this paper is testing the second and third implication. For that purpose, we need longitudinal matched worker firm data. Only by knowing the tenure distribution of the entire workforce of the firm, at all times, we can calculate the seniority of a worker. Though using this type of data has become more fashionable in recent years, they are still not widely available. We have been able to get access to such data on Denmark and Portugal. The two countries are a nice combination since their level of employment protection differs widely. We give a description of both data sets and the relevant institutions from these countries in the next subsection. Subsequently, we discuss the test of the three implications of our model, each in a separate subsection.

3.3.1 Data and relevant labour market institutions

For Denmark, we use the *Integrated Database for labor Market Research* (IDA) for 1980-2001, from the Danish Bureau of Statistics, which has been used previously e.g. in Mortensen (2003). IDA tracks every single individual between 15 and 74 years old. The labor market status of each person is recorded once a year, at November 30. The dataset contains a plant identifier, which allows the construction of the total workforce of a plant, and hence of the firm as a whole. There is information on earnings, occupation, education, and age, and on the plant's location, firm size, and industry. Industry is defined as the industry employing the largest share of the firm's workforce. Firm size is defined as the number of individuals holding primary jobs in that firm and earning a positive wage. The tenure of workers hired since 1980 can be calculated straightforwardly from the IDA. For workers hired between 1964 and 1980, the tenure can be calculated from a second dataset on the contribution histories to a mandatory pension program, the *ATP*. The tenure in job spells started before 1964 is left censored (less than 3% of the observations). We calculate potential experience as age-schooling-6.

For Portugal, we use the *Quadros de Pessoal* for 1991-2000 provided by the Ministry of Employment, which has been used before e.g. in Cabral and Mata (2003). It is based on a compulsory survey of firms, establishments and all their workers; the compulsory participation enhances the quality of the data. The information available is similar to that for Denmark except that workers' tenure is directly reported and the industry of the firm is that industry with the highest share of sales or, when the allocation by sales is not possible, the industry with the highest employment share. We use all full-time employees in their main job, aged between 16 and 66, and working for a firm located in Portugal's mainland. The hourly gross earnings were computed as the monthly base-wage plus seniority-indexed components plus other regularly paid components, divided by normal hours of work per month.

Table 3.1: Descriptive statistics Denmark and Portugal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Pooled sample												
DK 1980-2001	22364083	2771627	6870869	301015	164.22	5.41	20.25	12.05	37.79	0.35	0.63	4.68
	.	.	.	.	(59.94)	(5.58)	(12.19)	(3.05)	(11.95)	.	(0.72)	(2.44)
PT 1991-2000	11420191	3211990	4268149	330270	3.68	8.43	20.26	6.81	36.73	0.40	0.87	4.04
	.	.	.	.	(2.52)	(8.61)	(11.39)	(3.65)	(11.18)	.	(0.85)	(2.14)
Cross-section year 2000												
DK 2000	1087922	1087922	1087922	79188	23.53	5.26	21.10	12.69	38.98	0.36	0.61	4.86
	.	.	.	.	(9.02)	(5.78)	(11.68)	(2.74)	(11.75)	.	(0.73)	(2.48)
PT 2000	1403686	1403686	1403686	184957	4.06	7.93	20.47	7.59	37.14	0.42	0.75	3.64
	.	.	.	.	(2.75)	(8.54)	(11.42)	(3.83)	(11.08)	.	(0.81)	(2.18)

Columns' labels: (1) Observations, (2) Workers, (3) Spells, (4) Firms, (5) Average Real Hourly Wage (base year=2000, measure unit: Euro), (6) Average Tenure, (7) Average Potential Experience, (8) Average Education, (9) Average Age, (10) Proportion of Women, (11) Average Relative Log Rank, (12) Average Log Firm Size. (Standard deviations in paratheses.)

For both countries, we use data for all private sector jobs, except agriculture, fishing and mining. We eliminate outliers by deleting all wage observations lower than the legal minimum wage for each year and drop the top 1% of the wage distribution. Summary statistics for both countries are presented in Table 3.1, both for the pooled data and for 2000 separately[9]. There are several obvious differences between the two countries. The mean level of education is more than 5 years higher in Denmark than in Portugal, while the mean tenure is almost 3 years longer in Portugal than in Denmark. The number of firms is far higher in Portugal than in Denmark, and the average firm size in Portugal is only 30% of that in Denmark. Finally, Danes earn on average almost six times more than Portuguese.

Denmark and Portugal are both members of the European Union, both small open economies, and both with (in EU terms) low unemployment rates over recent years, see Nickell et al (2005). The two countries differ substantially with respect to their labour market flexibility and social safety net. In Denmark, wage bargaining is de facto done at the firm level over the observation period, although there are some collective industry-level bargaining agreements for minimum wages, see Kenworthy (2001). The Danish private labor market is characterized by very low EPL compared to most OECD countries, but similar to the United States, the United Kingdom and Australia, see OECD Employment Outlook (2004). The EPL applicable to privately owned firms is limited to basic provisions such as white-collar workers being given an advance notice and a minimum of EU enacted rules relating to mass layoffs. General rules and procedures for dismissal are absent, see also Albaek et al (1999). Unemployment benefits for wage earners are high and can be obtained for a long period, being generous compared to most other countries. In Portugal, wage negotiations start at the national level, defining a national minimum standard and setting guidelines for collective bargaining at a lower level. Massive collective agreements dominate the labor market as a result of extension mechanisms. However, individual firms are able to pay higher wages than those bargained at the aggregate level, see Cardoso and Portugal (2005). The EPL in Portugal is the other extreme compared to Denmark: according to the OECD Employment Outlook (2004), Portugal has about the highest overall summary index of all countries. The notice period for layoff is 60 days, and the severance pay for individual dismissals is 1 month per year of service, with a minimum indemnity of 3 months. The minimum duration of any contract is 6 months, although the law defines some exceptions. Following an unjustified dismissal the worker can be reinstated or compensated. The maximum duration of benefits in Portugal varies from 10

[9]Summary statistics for each separate industry (see Appendix B for our broad industry classification) are available upon request.

months for people aged under 25 and 30 months for those aged above 45. The replacement rate is 65 % of the previous wage, but the benefit cannot be below the minimum wage or 3 times above the minimum wage.

3.3.2 Testing Gibrat's law

Our theoretical model predicts log employment size of firm j, n_{jt}, to follow a random walk, apart from the dampening effect of the hiring and firing boundaries on short run fluctuations in n_{jt}. This regularity is known as Gibrat's law. Though slight deviations from Gibrat's law do not affect the main economic implications of the theoretical model, it is useful to have at least some idea how well this assumption is supported in the data. There is a massive literature on testing Gibrat's law, see e.g. De Wit (2005) or Sutton (1997). Here, we use two tests.

The first approach is laid out in Abowd and Card (1989) and Topel and Ward (1992) for log wages; we adapt this methodology for log firmsizes. First, we estimate

$$\Delta n_{jt} = \delta_0 + \delta_1 Z_{jt} + \varepsilon_{jt} \tag{3.10}$$

where Δ is the first difference operator and where Z_{jt} is vector of controls: age category of the firm, time effects and industry indicators. Second, we construct the autocovariance matrix of the residuals ε_{jt} of this regression. If n_{jt} follows a random walk, ε_{jt} should be uncorrelated across time t.

The resulting covariograms for (3.10) are reported in Table 3.2, both for the whole sample of firms and for the subsample of larger firms (at least 20 employees each year over the sample period of that firm). The evidence from Table 3.2. suggests that the process characterizing the residuals in firm size changes is very close to a random walk, except for the case of Denmark when including the small firms (column 1). In Portugal the residual auto-covariogram supports the specification of an autoregressive process very close to a unit root, although when using only the subsample of larger firms we find evidence of mild positive serial correlation in first differences.

The second approach follows the literature on testing for unit roots in panel data. Breitung-Meyer (1994) and Bond et al. (2005) show that for micro panels with large cross-sectional and small time dimension, OLS in levels is consistent and typically more efficient than more complex GMM and ML estimators. Consider a simple dynamic AR(1) panel data model, where for expositional brevity we do not include any covariates (in the estimation we use specifications where we control for age category of the firm, industry

Table 3.2: 1st Gibrat's law test: Residual autocovariances

	Denmark		Portugal	
Lag	(1)	(2)	(1)	(2)
0	0.1587	0.0424	0.1162	0.0255
	(0.0005)	(0.0112)	(0.0005)	(0.0007)
1	-0.0030	-0.00003	0.0002	-0.0001
	(0.0002)	(0.0005)	(0.0002)	(0.0003)
2	-0.0094	-0.0008	-0.0024	0.0012
	(0.0002)	(0.0003)	(0.0002)	(0.0004)
3	-0.0020	-0.0002	-0.0013	0.0006
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
4	-0.0016	-0.00004	-0.0008	0.0006
	(0.0002)	(0.0002)	(0.0003)	(0.0002)
5	-0.0008	-0.0006	-0.0010	0.0001
	(0.0002)	(0.0005)	(0.0003)	(0.0003)
6	-0.0008	-0.0002	-0.0013	0.0002
	(0.0002)	(0.0005)	(0.0004)	(0.0004)
N obs generating reg	1505926	79425	878919	66369

Specification (1) uses all the firms; specification (2) uses all firms that have at least 20 employees in each year of their life spans. All generating regressions use the first differenced log firm size as dependent variable and control for age of the firm, time and industry effects. (Robust standard errors in parentheses) and time effects):

$$n_{jt} = \beta n_{j,t-1} + u_{jt}, \tag{3.11}$$

where $u_{jt} \equiv (1-\beta)\gamma_j + v_{jt}$ and the initial firm size $n_{j1} = \alpha_0 + \alpha_1\gamma_j + \varepsilon_j\gamma_j$, with ν_{jt} and γ_j error terms such that E(γ_j) =E$(v_{jt}) = 0$ and E$(v_{jt}v_{js}) = 0$ for $t \neq s$. Mean stationarity in (3.11) requires $\beta < 1$, $\alpha_0 = 0$ and $\alpha_1 = 1$. In addition, covariance stationarity requires homoskedasticity over time of v_{jt}, i.e. Var$(v_{jt}) = \sigma^2_{v_j}$ and Var$(\varepsilon_j) = \sigma^2_{v_j}/(1-\beta^2)$. Bond et al (2005) show that under the null of $\beta = 1$ the OLS estimator of β in (3.11) is consistent. We refer to this estimator of β as the OLS estimator. Under the alternative $\beta < 1$, the OLS estimator is biased upwards; this is more the case when Var(γ_j)/Var(v_{jt}) is large. In cases where this difference in the variances is high, one could use the transformed statistic in Breitung and Meyer (1994), which estimates β from a transformed version of (3.11):

$$n_{jt} - n_{j1} = \beta(n_{j,t-1} - n_{j1}) + \varepsilon_{jt} \tag{3.12}$$

where $\varepsilon_{jt} = v_{jt} - (1-\beta)\left(n_{j1} - \gamma_j\right)$. The OLS estimator of (3.12) above is consistent again under the null and again upwards biased under the alternative $\beta < 1$, but this time the asymptotic bias does not depend on Var(γ_j)/Var(v_{jt}) when the process is mean stationary. The t tests based on these two estimators should be considered jointly when testing for the unit root and that tests based on other estimators that are consistent under both the null and under certain alternatives are found to have less power. For the purpose of our exercise, estimating β by least squares both in (3.11) and in (3.12) would provide a good indication whether the process is close to a random walk.

Table 3.3: 2nd Gibrat's law test: Unit root type regressions

	Denmark				Portugal			
	all firms		large firms		all firms		large firms	
Coef	OLS1	BM1	OLS2	BM2	OLS1	BM1	OLS2	BM2
β	.9361	.9208	.9755	.9806	.9594	.9537	.9791	1.043
	(.0003)	(.0006)	(.0012)	(.0030)	(.0004)	(.0009)	(.0011)	(.0030)
N obs	1505926		79425		878934		66340	
R^2	0.87	0.70	0.95	0.82	0.91	0.66	0.96	0.84
MSE	0.42	0.43	0.21	0.21	0.36	0.36	0.17	0.17

The dependent variable is logfirmsize in OLS columns and logfirmsize-initial logfirmsize in BM columns. Columns indexed 1 correspond to estimates using the sample of all firms, while columns 2 correspond to the sample of firms with at least 20 employees in each year of their life spans. Both regressions control for age of the firm, time and industry effects. (Robust standard errors in parentheses).

The results of the regressions in (3.11) and (3.12) are shown in Table 3.3, for the methodologies of both Bond et al. (2005) and Breitung and Meyer (1994), and both for the sample using all firms and the sample with only the large firms. The results are very similar to those in Table 3.2. Looking at the estimates for Denmark, the conclusions from the first Gibrat test are confirmed. While the coefficient on the first lag is somewhat lower than unity when using the sample of all firms, it approaches the unit root once we look at the subsample of large firms. The value for the MSE, a good estimate for the parameter σ of the theoretical model, is very similar for both countries; it is quite large (0.40) for the whole sample, but it is about half of that (0.20) for the subsample of larger firms, suggesting a lot of variation due to small firms. This confirms results from previous literature that Gibrat's law is more accurate for large firms, e.g. Jovanovic (1982). Similar conclusions hold for Portugal, where in the subsample of large firms we seem to be closer to the unit root. We conclude that the Gibrat law holds for large firms in both countries, while there is some mean reversion for small firms, in particular in Denmark.

3.3.3 Testing the LIFO separation rule

Next, we turn to the second prediction of our model, the Last-In-First-Out separation rule. Let the function $j(i,t)$ denote the firm j in which worker i is employed in period t. We drop the arguments of this function whenever the identification of the individual and the period of observation are clear. The seniority level q_{ijt} is defined as the log of number of workers employed at firm $j\,(i,t)$ at time t at least as long as or longer than worker i; thus, this number includes worker i herself. Hence, q_{ijt} is equal to n_{jt} at the moment t when worker i is hired (assuming that i is the only one hired at time t). Furthermore, for the most senior worker $q_{ijt} = 0$ because there is only one worker who is employed at the firm at least as long as herself. Then, the seniority index r_{ijt} is defined as the log of the ratio of the number of people employed at least as long as worker i to the size of firm j at time t, in logs

$$r_{ijt} \equiv n_{jt} - q_{ijt}. \tag{3.13}$$

The seniority index r_{ijt} is a reasonable proxy for the variable $z_t - q$, since z_t is equal to n_t, up to a constant, ηp, and except for the insulation of n_t from shocks in z_t when $p^- < p_t < p^+$, recall the setup of our theoretical model. Were the LIFO separation rule to apply literally, the seniority index r_{ijt} would be the only determinant of separation. However, there are two reasons why this is not likely to be the case. First, the workforce of the firm is not completely homogenous, so that a firm may wish to diminish its workforce in one skill class but not necessarily for other skill classes employed within that firm. This may disrupt a strict application of the LIFO separation rule. Second, workers separate not only due to shocks of the demand for the firm's product, but also due to worker specific shocks, e.g. when a worker's partner gets a new job in another city, which causes the worker to quit from his or her current job. A particularly important worker specific factor that does not fit in the LIFO model is retirement. Hence, our ambition is more limited than what would follow from a strict interpretation of the LIFO separation rule. We just want to show that r_{ijt} has a strong impact on the job separation rate.

We model the transition process by a mixed proportional hazard rates model with discrete time periods. This implies that the conditional probability of leaving the firm (i.e. the hazard rate) after T_{ijt} years of tenure can be written as:

$$\theta(r_{ijt}, Z_{ijt}, T_{ijt}, v_i) = \frac{\exp\left(\beta r_{ijt} + \gamma Z_{ijt} + \psi_{T_{ijt}} + v_i\right)}{1 + \exp\left(\beta r_{ijt} + \gamma Z_{ijt} + \psi_{T_{ijt}} + v_i\right)} \tag{3.14}$$

where Z_{ijt} is a vector of observed characteristics of the individual and the job, and where v_i represents the unobserved worker heterogeneity. We include a full set of dummies ψ_T for every tenure category, which is equivalent to a fully flexible specification of the baseline hazard. Identification of the parameter β of the seniority index r_{ijt} separate of the parameters of the baseline hazard ψ_T requires variation in r_{ijt} that is independent of the tenure T_{ijt}. Such independent variation is available since the seniority index also depends on the hiring and firing of other workers and hence is a non-deterministic function of tenure. A LIFO separation rule implies that β should be negative. For our estimation method we use a two mass-point distribution for the unobserved heterogeneity. We use up to 10 spells of an individual, which helps to estimate the unobserved heterogeneity distribution. The main reason for using a discrete time model is practical. For Denmark, the worker is observed only once per year. Hence, we cannot observe the exact moment at which the worker enters or leaves the firm.[10] In addition, short spells are underrepresented since a worker has to stay at least till the next period of observation. With the data at hand, we cannot correct for these problems in a continuous time model and even though we do not claim that a discrete time analysis solves all these problems, this is the simplest accurate representation of the data.

As noted before, older workers may leave the firm for retirement. This process is likely to run counter to the LIFO separation rule, since retired workers tend to have long tenures. Therefore, we exclude workers above the age of 55 from the analysis. Spells started before

[10] For Portugal, tenure is reported in months; we use this information in the estimation. For the rest, the modelling is identical to that for Denmark.

the age of 55 and finished afterwards are thus right censored. Women are also more likely to leave the firm for non-participation. Hence, we separate our results for men and women. We delete spells that are left censored since we cannot compute the seniority of an individual for the periods before she enters our observation sample. Since seniority affects the probability that the individual survives till the start of the sample period, we cannot easily correct for left censoring. Deleting the left-censored spells implies that we have a maximum of 22 years of tenure in Denmark and 10 for Portugal. The vector Z_{ijt} includes education, potential experience and indicators for region, industry and occupation.

Table 3.4 lists the main results. We find a negative and significant impact of seniority for both women and men, with small differences between these categories, in both Denmark and Portugal, in accordance with the LIFO separation rule. Furthermore, education and experience have a negative impact on the job separation hazard. Though the actual coefficients are not reported here, we also find negative duration dependence and evidence of unobserved heterogeneity, in both countries.[11] Apparently, seniority does not pick up all the variation in separation rates over the course of a job spell. There are two explanations for this phenomenon. First, as noted before, our seniority index might not exactly correspond to the actual layoff ordering, since the firm's workforce is likely to be heterogeneous, with separate LIFO ordering applying to subsets of the workforce. This is equivalent to measurement error in our seniority index r_{ijt}, leading to an attenuation bias in the estimate of β and unobserved variation in the seniority index being picked up by correlated variables, i.e. the tenure dummies ψ_T. Second, not all separations are driven by the fluctuations in the demand for the firm's product, and hence, the log seniority index. In particular, some separations are driven by the worker and the firm learning about the quality of the match. This learning process leads to a hump shape hazard, with many separations early on and a quickly declining hazard for higher elapsed tenures, see e.g. Jovanovic (1979a).

The analysis in Section 2.8 shows that firing cost increases the expected duration of job spells. Since EPL is more stringent in Portugal than in Denmark, one would expect longer job durations in Portugal. In Figure 2 below, we present the estimated cumulative distribution of completed tenures for a job spell of a male with 12 years of education, which started at the age of 25, and having $r_{ijt} = 0$ along the whole spell, for both Denmark and Portugal. Indeed, the expected tenure is clearly higher in Portugal than in Denmark. This conclusion generalizes to any other type of worker.

[11] The full estimation results are available upon request.

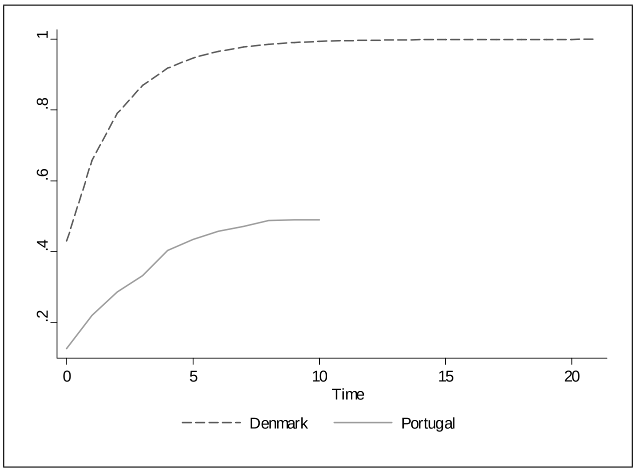

Figure 3.2: CDF expected completed tenure Denmark and Portugal

Table 3.4: Main results LIFO test

	Denmark		Portugal	
	Males	Females	Males	Females
Logrank	-0.0577	-0.0357	-0.0549	-0.0669
	(0.0019)	(0.0025)	(0.0054)	(0.0065)
Education	-0.1169	-0.1267	-0.1204	-0.1446
	(0.0003)	(0.0005)	(0.0009)	(0.0012)
Experience	-0.0771	-0.0732	-0.0490	-0.0656
	(0.0001)	(0.0001)	(0.0003)	(0.0004)
N obs	10788368	5990891	2118405	1488687

The estimation also controls for occupation, region and industry indicators. (Standard errors in parantheses)

3.3.4 Testing dependency of wages on seniority

The third and main prediction of the model is the dependency of wages on seniority. This can be tested by extending the standard specification of the log earnings equation with the seniority index, r_{ijt}, as defined in (3.13) above. Consider the following specification of log wages w_{ijt}

$$w_{ijt} = \alpha + \chi X_{ijt} + \gamma T_{ijt} + \delta r_{ijt} + \zeta n_{jt} + \varepsilon_{ijt}, \tag{3.15}$$

where X_{ijt} is experience. We omit higher order terms in experience and tenure and other controls (including time effects) from equation (3.15) for the sake of convenience, but include them in the estimation. The unobservable term can be decomposed into four orthogonal components, a match, a firm, a worker, and an idiosyncratic effect[12]

$$\varepsilon_{ijt} = \varphi_{ij} + \psi_j + \mu_i + \nu_{ijt}. \tag{3.16}$$

The idiosyncratic effect ν_{ijt} can also include measurement error. There are all kinds of reasons for ϕ_{ij}, ψ_j, and μ_i to be correlated to T_{ijt}, see Topel (1991) or Altonji and Williams (2005): good worker-firm relationships tend to survive, as the worker and the firm learn about the quality of their match and bad matches are broken up, leading to a positive correlation between $\varphi_{ij} + \psi_j + \mu_i$ and T_{ijt}. Search theories imply that workers sample new jobs from a job offer distribution. The longer this selection process is going on, the higher the expected value of $\varphi_{ij} + \psi_j$ since bad jobs do not survive, leading to a positive correlation between $\varphi_{ij} + \psi_j$ and T_{ijt}. There are two obvious solutions to this problem, either within-job first differencing (FD) or adding fixed effects for every job spell (FE). First differencing yields

$$\Delta w_{ijt} = \chi + \gamma + \delta \Delta r_{ijt} + \zeta \Delta n_{jt} + \Delta \nu_{ijt}. \tag{3.17}$$

Adding fixed effects per job spell is equivalent to estimating (3.15) by taking deviations from the mean over time, within a job spell:

$$\widetilde{w}_{ijt} = (\chi + \gamma) \widetilde{T}_{ijt} + \delta \widetilde{r}_{ijt} + \zeta \widetilde{n}_{jt} + \widetilde{\nu}_{ijt}, \tag{3.18}$$

where the upper tilde denotes deviations from the mean per job spell, e.g. $\widetilde{w}_{ijt} = w_{ijt} - \overline{w}_{ijt}$, with $\overline{w}_{ijt}$ the mean over time of w_{ijt}. We exclude $\widetilde{X}_{ijt}$ from (3.18) because it is perfectly collinear with $\widetilde{T}_{ijt}$. In both specifications above, it is immediately clear that the first order effects of tenure and experience are not separately identified. However, this problem does not affect the estimation of δ, since r_{ijt} is not perfectly correlated to T_{ijt}. The choice between the FE and FD estimators above depends on the error structure of v_{ijt}. The closer is v_{ijt} to a unit root, the more efficient is the FD method; the closer v_{ijt} is to being serially uncorrelated, the more efficient estimation method is the FE estimator. Previous empirical studies have typically found a high degree of autocorrelation in v_{ijt}, even close to a unit root, see for instance Abowd and Card (1979) and Topel and Ward (1992). From that perspective, equation (3.17) is likely to be most efficient. However, this equation assumes that the effect of r_{ijt} and n_{jt} on w_{ijt} is immediate. Any lagged impact will not be captured after first differencing. From that perspective, equation (3.18) is preferred,

[12] This formulation is similar to Topel (1991: 150), except that we add a firm effect and that we delete the subscript t from the match effect ϕ_{ij}, as Topel does in his application.

since there lagged effects of r_{ijt} and n_{jt} will be captured. Hence, one would expect higher estimates for δ and ζ from using equation (3.18) than from (3.17).[13] In the strict version of our model, where separation is completely governed by the LIFO separation rule, r_{ijt} and n_{jt} are perfectly correlated within a job spell, since more senior workers will never leave the firm before worker i, so that the only variation in r_{ijt} comes from variation in n_{jt}. The same argument applies to $\widetilde{r}_{ijt}$ and $\widetilde{n}_{jt}$. Hence, δ and ζ are not separately identified in that world neither in equation (3.17), nor in (3.18). Happily, LIFO does not apply in a strict sense. The most compelling reason for a violation of the LIFO separation rule is workers' retirement, but also other individual specific shocks discussed earlier in this section. These separations allow separate identification of δ and ζ with FE and FD estimators.

Table 3.5: Residual autocovariances for within-job logwage innovations

Lag	Denmark	Portugal
0	0.0231	0.0273
	(0.00002)	(0.00007)
1	-0.0043	-0.0082
	(0.00001)	(0.00006)
2	-0.0006	-0.0008
	(8.7e-06)	(0.00003)
3	-0.0003	-0.0004
	(9.0e-06)	(0.00003)
4	-0.0003	9.2e-06
	(9.5e-06)	(0.00003)
5	-0.00008	-0.00008
	(0.00001)	(0.00004)
6	-0.0001	-0.0006
	(0.00001)	(0.00005)
N obs generating reg	14907897	5758655

The generating regressions are the FD wage regressions with logrank includes, see the FD2 columns in the next table. (Robust standard errors in parentheses)

First, we check the characteristics of the dynamic process of v_{ijt}. Table 3.5 reports the variance-covariance of Δv_{ijt}, analogous to what we did for log firm sizes in Table 3.2. For both countries, the covariance of ε_{ijt} with its first lag is substantial, the covariance with higher lags is negligible. Hence, this is well approximated by an MA(1) process, made up of a mixture of permanent and transitory shocks. Abowd and Card (1979) and Topel

[13] We report robust standard errors, so that correlation between the residuals over time does not affect their validity.

and Ward (1992) find similar results for the United States. The standard deviation of the permanent shocks can be calculated as 0.12 for Denmark and 0.10 for Portugal.[14] These numbers are of the same order of magnitude as found for the United States.

[14] Let q_{ijt} and u_{ijt} be the transitory and permanent shock respectively. Then:

$$\Delta v_{ijt} = u_{ijt} + q_{ijt} - q_{ij,t-1}.$$

Hence: $\text{Var}(\Delta v_{ijt}) = \text{Var}(u_{ijt}) + 2\text{Var}(q_{ijt})$ and $\text{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1}) = -\text{Var}(q_{ijt})$, so that: $\text{Var}(u_{ijt}) = \text{Var}(\Delta v_{ijt}) + 2\text{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1})$.

Table 3.6: FE and FD Wage regressions for the entire private sector in Denmark and Portugal

	Denmark				Portugal			
	FD1	FD2	FE1	FE2	FD1	FD2	FE1	FE2
logrank		.003***		.008***		.016***		.022***
		(.0003)		(.0003)		(.0005)		(.0005)
logfsize	.013***	.011***	.026***	.021***	.025***	.015***	.040***	.028***
	(.0002)	(.0003)	(.0003)	(.0003)	(.0004)	(.0005)	(.0004)	(.0004)
tenure+exper	.047***	.045***	.010***	.007***	.068***	.065***	.059***	.055***
	(.0003)	(.0003)	(.0001)	(.0002)	(.0005)	(.0005)	(.0003)	(.0003)
tenure^2	.191***	.199***	-.052***	-.036***	-.086***	-.069***	-.083***	-.067***
	(.002)	(.002)	(.002)	(.002)	(.003)	(.003)	(.002)	(.002)
tenure^3	-.101***	-.105***	.014***	.008***	.027***	.021***	.024***	.019***
	(.001)	(.001)	(.099)	(9.88e-07)	(.001)	(.001)	(.0007)	(.0007)
tenure^4	.002***	.002***	-.0009***	-.0002	-.003***	-.002***	-.003***	-.002***
	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)	(.00009)	(.00009)
exper^2	-.224***	-.223***	.099***	.100***	-.204***	-.204***	-.149***	-.147***
	(.002)	(.002)	(.0006)	(.0006)	(.004)	(.004)	(.002)	(.002)
exper^3	.039***	.039***	-.039***	-.039***	.043***	.043***	.030***	.029***
	(.0007)	(.0007)	(.0002)	(.0002)	(.001)	(.001)	(.0007)	(.0007)
exper^4	-.003***	-.003***	.004***	.004***	-.003***	-.003***	-.002***	-.002***
	(.00007)	(.00007)	(.00003)	(00003)	(.0001)	(.0001)	(.00007)	(.00007)
N obs	14907897		22364083		5758655		10743244	
Workers	2116307		277162		1752000		3092329	
Spells	3745050		6870869		1965560		4053649	
Firms	221106		301015		206361		322502	

The dependent variable is the (Δ) time-detrended log real hourly wage for the (FD) FE columns; the covariates have Δ in front for FD columns. Columns 1 report results for the same regressions as corresponding columns 2, but without logrank included as covariate. The higher order polynomials in tenure and experience are divided by the corresponding powers of 10. All regressions also control for region, industry and occupation indicators. Significance levels: * : 10% ** : 5% *** : 1%. (Robust standard errors in parentheses).

This evidence suggests that in terms of efficiency of the estimation method we might prefer FD, while in terms of allowing for a lagged effect of Δr_{ijt} on Δw_{ijt} we might prefer FE. Hence, we report both the FD and FE estimators. Our regressions control for up to quartic terms in tenure and experience, log firmsize and industry, occupation, and region dummies. In Table 3.6 we report the results[15]. We present estimation results for two specifications, one excluding log seniority r_{ijt} and another including it. We can draw the following conclusions. First, all coefficients for log seniority are positive and statistically significant. Second, the coefficients are larger for FE than for FD, as was expected, because FE allows for a lagged effect of r_{ijt} on w_{ijt}, while FD does not. Third, comparing the estimation results with and without seniority, including seniority reduces the estimates for the first order effect of tenure + experience and for log firmsize by 5-30 %. The coefficients for the higher order effects of tenure and experience are hardly affected by including seniority. The effect of log firmsize and tenure on wages is at least partly a proxy for the effect of seniority. Of the three variables, tenure, firm size, and seniority, we can expect seniority to be measured with the greatest amount of measurement error. Apart from straightforward reporting errors, the main source of measurement error in tenure is who exactly is the relevant employer. Some job changes might either be classified as between firms, justifying the tenure clock being set back to zero, or as within firms, which does not affect the tenure clock. However, this source of measurement error only affects changes at the borderline of the definition of a firm. This is likely to be only a small fraction of the firm's workforce. However, misclassification of the tenure of even a single worker can affect the measurement of the seniority of all other workers of the firm. In general, any measurement error in tenure or firm size automatically feeds into seniority, while on top of that, seniority is also affected by measurement errors because separate seniority statistics are likely to apply for subgroups of the workforce. Both the upward effect on the coefficients for tenure and log firm size and downward effect on the coefficient for seniority of the measurement error in seniority can therefore be expected to be substantial. Finally, the effect of seniority is twice as high in Portugal than in Denmark.[16] It is tempting to link this difference between Denmark and Portugal to the differences in EPL in both countries, but our theoretical model does not allow for a link

[15] Results for the same analysis performed for each of the broad industry categories defined in Appendix B are qualitatively identical (with some heterogeneity in the magnitude of the estimated seniority coefficients, among the various industries) to the results at the national level in Table 3.6. They are available upon request from the authors.

[16] A corrolary of that is that also the return to tenure is very low in Denmark, compared to most other countries, and certainly so when compared to Portugal. The fact that the common linear effect of tenure and experience is further reduced when accounting for seniority rank suggests however that seniority rank, keeping in mind the proportions, is a much more important factor in wage determination than tenure (whose still significant impact, is in our interpretation due to measurement error in seniority).

between the bargaining power β and the firing cost W.

Returns to seniority within gender and education subgroups

We repeat the analysis separately for males and females, and for low- and high-educated workers. The results are reported in Table 3.7. The results for male and female categories do not differ much. The only apparent exception is for Denmark, when using the FE estimator, where the estimated coefficient for males is twice as high as for females, though they are the same when using the FD estimator. At the same time, and linked to the previous observation, the estimates by gender categories really do not differ much from the estimate when using the whole samples, for either country. Our interpretation is therefore that seniority within gender categories is not more relevant for wage determination than seniority within the firm as a whole and hence, splitting by gender is not likely to attenuate the measurement error in seniority index.

The estimation results for education groups show that the effect of seniority is much larger for high educated workers than for low educated workers. The impact of seniority on wages is lower for the low-educated workers, compared to corresponding estimates from Table 3.6, and the FE estimate is even significantly negative for Denmark, though small in absolute value. The impact of seniority on wages within the high educated group is much larger, both in Denmark and in Portugal. These results are consistent with the fact that high educated workers have steeper wage-tenure profiles than their low-educated peers. At the same time, they give support to the fact that the relevant seniority hierarchy within the firm is already more realistically captured when accounting for education levels.

Table 3.7: FE and FD Regressions by gender and education rank groups

	Denmark				Portugal			
	Gender categories							
	Females		Males		Females		Males	
	FD	FE	FD	FE	FD	FE	FD	FE
logrank	.005***	.005***	.005***	.010***	.015***	.019***	.014***	.019***
	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0007)	(0.0006)	(0.0007)	(0.0006)
logfsize	.002***	.014***	.014***	.025***	.015***	.028***	.019***	.031***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0007)	(0.0006)	(0.0006)	(0.0006)
ten+exp	.032***	.009***	.052***	.007***	.053***	.042***	.080***	.073***
	(0.0004)	(0.0002)	(0.0004)	(0.0002)	(0.0007)	(0.0005)	(0.0007)	(0.0005)
N obs	5049388	7745676	9858509	14618407	2300767	4353808	3457888	6389436
	Education categories							
	HighEduc		LowEduc		HighEduc		LowEduc	
	FD	FE	FD	FE	FD	FE	FD	FE
logrank	.010***	.020***	.002***	-.002***	.029***	.032***	.013***	.016***
	(.0003)	(.0004)	(.0004)	(.0004)	(.002)	(.002)	(.0005)	(.0005)
logfsize	.007***	.016***	.014***	.024***	.026***	.026***	.016***	.031***
	(.0004)	(.0004)	(.0005)	(.0005)	(.002)	(.002)	(.0005)	(.0004)
(ten+exp)	.040***	.006***	.031***	.006***	.116***	.099***	.056***	.049***
	(.0004)	(.0002)	(.0007)	(.0002)	(.002)	(.001)	(.0005)	(.0003)
N obs	9567345	14054988	5268672	8309095	259793	536920	5492034	10206324

The dependent variable is the (Δ) time-detrended log real hourly wage for the (FD) FE columns; the covariates have Δ in front for FD columns. Logrank has been computed separately for each category. "LowEduc" stands for the category of people with at most 12 years of education. All regressions include also up to 4th order polynomials in tenure and experience and indicators for region, occupation and industry. Significance levels: $*$: 10% $**$: 5% $***$: 1%. (Robust standard errors in parentheses).

Returns to seniority and firm monopoly power

Our theoretical model also predicts that the return to seniority, β/η, is partly driven by the degree of monopoly power, η^{-1}. We test this hypothesis by analyzing whether the variation in the return to seniority across industries can be explained by the degree of monopoly power in each industry. We take the log of the number of firms in each industry as proxy for the degree of monopoly power. We regress the estimated coefficient for seniority for each industry on this log number of firms and a constant term, using both simple OLS and Weighted Least Squares (WLS) specifications. We use two measures as "number of firms" in an industry: the sum of all firms that were at any time active in that industry during the sample period, and respectively, the median number of firms over the sample period. We use two industries classifications, a broad classification with 12 industries for Denmark and 11 for Portugal, see Appendix B, and a more refined classification where we use all 2-digit Standard Industry Classification (SIC) sub-categories available, increasing the number of observations in the regressions to 40 for Denmark and 49 for Portugal. For our prediction to be verified, we expect negative estimates of the coefficients of log number of firms.

The estimation results for the regressions of returns to seniority on the log number of firms by industry are presented in Table 3.8. Most of the estimated coefficients of interest are not significantly different from 0 (though most are slightly negative in magnitude), both when using the WLS and the OLS methods and regardless of using as dependent variables the FD or the FE coefficients previously estimated in this paper, and as independent variables the sum or the median of the number firms in an industry. There are very few cases where the results are statistically significant: when using the broad industry categories for Portugal we get significant coefficients of the expected sign with the FE method, but significant coefficients of the opposite sign in Denmark; when using the OLS for 2-digit industries in Portugal we get significantly positive coefficients for the FD method and again significantly positive when using the WLS for the FD, sum of firms, and FE, median of firms. In conclusion, we regard this test as inconclusive. The explanatory variables used as proxy for the monopoly power of an industry are not strong enough to isolate the effect of the degree of monopoly power on the return to seniority.

Table 3.8: Monopoly power test

	Denmark				Portugal			
	Sum		Median		Sum		Median	
	FD	FE	FD	FE	FD	FE	FD	FE
OLS Broad industry categories								
log Nfirms	-0.00009	0.0058**	0.0003	0.0060**	0.0046	-0.0069*	0.0046	-0.0068*
	(0.0014)	(0.0024)	(0.0014)	(0.0022)	(0.0028)	(0.0032)	(0.0027)	(0.0030)
constant	0.0107	-0.0435*	0.0072	-0.0378*	-0.0275	0.0892**	-0.0234	0.0817**
	(0.0135)	(0.0227)	(0.0111)	(0.0178)	(0.0267)	(0.0303)	(0.0230)	(0.0262)
WLS Broad industry categories								
log Nfirms	-0.0008	0.0053**	-0.0004	0.0002	-0.0004	-0.0057	-0.0002	0.0032
	(0.0014)	(0.0024)	(0.0014)	(0.0013)	(0.0031)	(0.0035)	(0.0030)	(0.0029)
constant	0.0174	-0.0395	0.0128	0.0075	0.0229	0.0778**	0.0212	-0.0102
	(0.0134)	(0.0227)	(0.0117)	(0.0109)	(0.0302)	(0.0340)	(0.0269)	(0.0251)
OLS 2-digit industry categories								
log Nfirms	-0.0010	-0.0010	-0.0010	-0.0007	0.0136**	-0.0003	0.0112**	-0.0011
	(0.0011)	(0.0016)	(0.0012)	(0.0016)	(0.0058)	(0.0030)	(0.0056)	(0.0028)
constant	0.0213**	0.0224*	0.0198**	0.0193*	-0.0852*	0.0314	-0.0559	0.0358*
	(0.0093)	(0.0125)	(0.0080)	(0.0108)	(0.0440)	(0.0228)	(0.0367)	(0.0187)
WLS 2-digit industry categories								
log Nfirms	-0.0012	0.0013	-0.0013	-0.0010	0.0094*	-0.0004	0.0075	0.0091*
	(0.0013)	(0.0015)	(0.0013)	(0.0013)	(0.0048)	(0.0025)	(0.0046)	(0.0047)
constant	0.0227**	-0.0022	0.0213**	0.0194**	-0.0515	0.0306	-0.0294	-0.0421
	(0.0110)	(0.0125)	(0.0085)	(0.0091)	(0.0365)	(0.0195)	(0.0307)	(0.0315)

The dependent variable is the estimated coefficient of seniority rank from FD and respectively FE regressions, by industry. "Nfirms" is measured as the sum of all firms (corresponding to columns labeled "Sum") or as the median of firms (corresponding to columns "Median"), over the sample period, respectively. There are 12 'broad industries' in Denmark and 11 in Portugal, while the number of 2-digit SIC sub-categories is 40 in Denmark and 49 in Portugal. Significance levels: $*$: 10% $**$: 5% $***$: 1%. (Standard errors in parentheses).

3.4 Summary and conclusions

What have we established beyond reasonable doubt in this paper? We have shown that for Denmark and Portugal part of what has been known as the wage return to tenure is in fact a return to *seniority*, that is, the position of the worker in the tenure hierarchy of her firm. This implies that standard explanations of the return to tenure, like Jovanovic's learning or the classic search models, and subsequent versions of these, cannot provide the full story, if only because these explanations focus solely on the features of the worker herself (in case of learning, her ability; in case of search, her job offer history), while the return to seniority links the fate of the worker to that of the firm as a whole. A return to seniority implies that a worker is to some extent shareholder in her own firm. Hence, it makes the link between labor economics and finance.

Our theoretical model provides a special interpretation of the return to seniority, as being due to a hold up problem, where firms pay the full cost of the specific investment, while workers capture part of the return. This setup leads to inefficiently low hiring. All these conclusions are conditional on the assumption that the firm bears the full cost of specific investments, an assumption that has not been tested empirically in this paper. How to do that remains an open question. An indirect answer can be obtained by analysing who is queueing for whom: when workers queue for jobs, so that there is unemployment, firms are held up by their incumbent workforce; when it is the other way around, and there are vacancies, workers are held up by their employer. As long as workers are risk neutral and either investment or wages are contractible, efficient hiring can be obtained by using the sharing rule of the costs for the one, to mirror the sharing rule for the other, thereby satisfying the Hosios (1990) condition. When workers are risk averse, efficiency can only be obtained when both investment and wages are contractible, such that the costs of investment are fully attributed to the firm and there is no seniority profile. Any other allocation assigns part of the risky return to the risk averse player. In that sense, our estimation results point to incompleteness in the insurance market. Nevertheless, our analysis does not imply that LIFO layoff rules are bad per se. They can offer a useful protection to the property rights of incumbent workers on their share of the specific investment, thereby helping the firm to solve a commitment problem. Without a resolution of this commitment problem, incumbents would have all reasons not to cooperate in the transfer of tacit knowledge to newly hired workers.

We have established the existence of a return to seniority for Denmark and Portugal. Whether such a return exists in other countries, in particular in the United States, remains an open question. We bet it does; the large return to tenure in the United States as compared to Denmark and Portugal strongly suggests so. One might argue that returns to seniority are largely driven by legal institutions, and that these institutions are entirely different and more market oriented in the United States. We think however that the economic mechanisms for having a LIFO layoff rule exist everywhere, and that the legal institutions might very well just be a formalisation of rules of conduct and implicit contracts that would have emerged anyway.

The return to seniority is twice as high in Portugal than in Denmark. It is tempting to relate this difference to the much more extensive Employment Protection Legislation (EPL) in the Portugal. Nevertheless, this does not follow from our theoretical model. Compared to, for example, Bentolila and Bertola (1990), our analysis has the advantage that it allows for the effect of EPL on wages, but this does not imply a higher return to seniority. What would be an interesting extension of our analysis is to allow for the fact that empirically EPL goes up with tenure. Till sofar, including this feature in theoretical models was cumbersome from an analytical point of view, but in the framework presented here this is likely to be doable. With an eye on the missing market for elderly workers in many European countries, this seems to be a worthwhile extension. We leave this for future research.

Our model suggests that hold up problems reduce turnover, and thereby specific investment (because turnover requires new specific investment to be made). This conclusion is contingent on the way specific investment is modelled here, namely as a fixed amount to be invested in a one-shot at the beginning of the job. When the amount of investment can vary both in size and in timing, this conclusion might change. Then, a longer expected job duration might invoke more specific investment, which in turn would lengthen the expected job duration since the productivity at the job is raised relative to the productivity at the outside market. In such a world, a firm responds along two margins of adjustment, when the demand for its product goes up. First, it will hire additional workers, and second, it will expand the specific investment in its incumbent workforce. This model would provide further legitimation for a LIFO rule, not as legal constraint, but as an efficient economic institution. Again, we postpone this for future research.

Appendices Chapter 3: Derivation. Industry categories

3.A Derivation solution non-linear system

Equation (3.6) can be written as

$$\begin{aligned} 0 &= \left(\mu^- - \beta\right) R - \psi + \left(\mu^- - \beta\right) C^- - \left(\mu^- - \beta\right) C^+, \qquad (3.19) \\ 0 &= \left(\mu^- - \beta\right) R - \psi\beta + \left(\mu^- - \beta\right) \mu^- C^- - \left(\mu^- - \beta\right) \mu^+ C^+, \\ \left(\mu^- - \beta\right) I &= \left(\mu^- - \beta\right) ER - \psi E_\beta + \left(\mu^- - \beta\right) E^- C^- - \left(\mu^- - \beta\right) E^+ C^+, \\ 0 &= \left(\mu^- - \beta\right) ER - \psi\beta E_\beta + \left(\mu^- - \beta\right) \mu^- E^- C^- - \left(\mu^- - \beta\right) \mu^+ E^+ C^+, \end{aligned}$$

where

$$\begin{aligned} \eta\lambda^- &\equiv \mu^- < 0, \eta\lambda^+ \equiv \mu^+ > 1, \psi \equiv \frac{\mu^-}{\rho} < 0, \Delta \equiv p^+ - p^- \\ C^- &\equiv B^- \exp\left[\mu^- p^-\right], C^+ \equiv B^+ \exp\left[\mu^+ p^-\right], R \equiv r\left(\eta^{-1}\right)^{-1} \exp\left[p^- - \pi\right] > 0, \\ E &\equiv \exp\left[\Delta\right], E_\beta \equiv \exp\left[\beta\Delta\right], E^- \equiv \exp\left[\mu^-\Delta\right], E^+ \equiv \exp\left[\mu^+\Delta\right]. \end{aligned}$$

Elimination of C^- yields:

$$\begin{aligned} 0 &= \left(\mu^- - 1\right) R - \psi - \left(\mu^- - \mu^+\right) C^+, \qquad (3.20) \\ \left(\mu^- - \beta\right) I &= \left(\mu^- - \beta\right)\left(E - E^-\right) R - \psi\left(E_\beta - E^-\right) - \left(\mu^- - \beta\right)\left(E^+ - E^-\right) C^+, \\ 0 &= \left(\mu^- - \beta\right)\left(E - \mu^- E^-\right) R - \psi\left(\beta E_\beta - \mu^- E^-\right) - \left(\mu^- - \beta\right)\left(\mu^+ E^+ - \mu^- E^-\right) C^+. \end{aligned}$$

Rewriting the first equation yields equation (3.7). The system of equations (3.20) can be rewritten as a system of linear equation in R, C^+, and $\left(\mu^- - \beta\right) I$

$$\begin{aligned} \begin{bmatrix} R \\ C^+ \\ \left(\mu^- - \beta\right) I \end{bmatrix} &= \psi \begin{bmatrix} \mu^- - 1 & -\left(\mu^- - \mu^+\right) & 0 \\ \left(\mu^- - \beta\right) D_{0-} & -\left(\mu^- - \beta\right) D_{+-} & -1 \\ \left(\mu^- - \beta\right) D_{10\mu-} & -\left(\mu^- - \beta\right) D_{\mu+\mu-} & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ D_{\beta-} \\ D_{\beta\beta\mu-} \end{bmatrix} \qquad (3.21) \\ D_{0-} &\equiv E - E^-, D_{+-} \equiv E^+ - E^-, D_{\beta-} \equiv E_\beta - E^-, \\ D_{10\mu-} &\equiv E - \mu^- E^-, D_{\mu+\mu-} \equiv \mu^+ E^+ - \mu^- E^-, D_{\beta\beta\mu-} \equiv \beta E_\beta - \mu^- E^-. \end{aligned}$$

Since $R > 0$, the first equation of the solution to equation (3.21) implies that

$$\left(\mu^+ - \mu^-\right) D_{10\mu-} - \left(1 - \mu^-\right) D_{\mu+\mu-} = -\mu^+\left(E^+ - E\right) + \mu^+\mu^-\left(E^+ - E^-\right) - \mu^-\left(E - E^-\right) < 0$$

Hence, Δ and C^+ should be positive for a solution to exist. The third equation of this system reads

$$I = \frac{\psi}{\mu^- - \beta}\left[\frac{(\mu^- - \beta)(D_{0-}D_{\mu+\mu-} - D_{+-}D_{10\mu-})}{(\mu^+ - \mu^-)D_{10\mu-} - (1-\mu^-)D_{\mu+\mu-}} + \frac{(\mu^+ - \mu^-)D_{0-} - (1-\mu^-)D_{+-}}{(\mu^+ - \mu^-)D_{10\mu-} - (1-\mu^-)D_{\mu+\mu-}}D_{\beta\beta\mu-} - D_{\beta-}\right] \tag{3.22}$$

which is an implicit equation in Δ. Since

$$\begin{aligned} I(0,\beta) &= 0, I(\infty,\beta) = \infty, \\ I_\Delta(\Delta,\beta) &> 0, I_\beta(\Delta,\beta) < 0, \end{aligned}$$

a unique positive solution for Δ exist for every $I > 0$, and $d\Delta/d\beta$ is positive.

With firing cost, the first equation of (3.19) reads

$$-(\mu^- - \beta)W = (\mu^- - \beta)R - \psi + (\mu^- - \beta)C^- - (\mu^- - \beta)C^+.$$

Elimination of C^- yields

$$-\mu^- W = (\mu^- - 1)R - \psi - (\mu^- - \mu^+)C^+.$$

This can be written as

$$\begin{bmatrix} R \\ C^+ \\ (\mu^- - \beta)I \end{bmatrix} = \psi \begin{bmatrix} \mu^- - 1 & -(\mu^- - \mu^+) & 0 \\ (\mu^- - \beta)D_{0-} & -(\mu^- - \beta)D_{+-} & -1 \\ (\mu^- - \beta)D_{10\mu-} & -(\mu^- - \beta)D_{\mu+\mu-} & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 - \rho W \\ D_{\beta-} \\ D_{\beta\beta\mu-} \end{bmatrix}$$

All conclusions in the text follow from this equation.

3.B Broad industry categories

1. Manufacturing
2. Electricity, gas and water supply
3. Construction
4. Wholesale and retail trade; repairs
5. Hotels and restaurants
6. Transport, post and communications
7. Financial intermediation
8. Real estate, renting and business activities

9. Public administration and defense; compulsory social security
10. Education
11. Health and social work
12. Other community, social and personal service activities

Note: For Portugal we miss category 9 (no firms are privately owned in that sector).

Chapter 4

A Social Network Analysis of Occupational Segregation

4.1 Introduction

Occupational segregation between various social groups is an enduring and pervasive phenomenon, with important implications for the labor market. Richard Posner recently pointed out that "a glance of the composition of different occupations shows that in many of them, particularly racial, ethnic, and religious groups, along with one or the other sex and even groups defined by sexual orientation (heterosexual vs. homosexual), are disproportionately present or absent"[1]. There are countless empirical studies within sociology and economics that document the extent of occupational segregation. Most studies investigating the causes of occupational segregation agree that 'classical' theories such as taste or statistical discrimination by employers cannot alone explain occupational disparities and their remarkable persistence. While several meritorious alternative theories were to date considered, scientists with long-standing interest in the area, such as Kenneth Arrow (1998), particularly referred to modeling the social network interactions as a very

[0]This chapter is based on a revised version of Buhai and Van der Leij (2006). We are grateful to Willemien Kets for carefully reading a previous version and providing us with valuable comments. For very useful discussion and suggestions, at different stages of this paper, we also thank Michèle Belot, Sanjeev Goyal, Maarten Janssen, Joan de Martí, Friederike Mengel, James Montgomery, José Luis Moraga-González, Paolo Pin, Gilles Saint-Paul, Ott Toomet, Jan van Ours, Fernando Vega-Redondo, Yves Zenou and audiences in seminars and conferences at University College London, Tinbergen Institute Amsterdam, Tinbergen Institute Rotterdam, Tartu University, NAKE in Utrecht, WEHIA in Essex, SMYE in Geneva, EEA in Vienna and CTN in Venice. The usual disclaimers apply.

[1]The quote is from a post in "The Becker-Posner Blog", see http://www.becker-posner-blog.com. Posner goes on by giving a clear-cut example of gender occupational segregation: "a much higher percentage of biologists than of physicists are women, and at least one branch of biology, primatology, appears to be dominated by female scientists. It seems unlikely that all sex-related differences in occupational choice are due to discrimination"

promising avenue for further research in this context.

In this paper we consider therefore a simple social interactions model in order to investigate the network channel leading to occupational segregation and wage inequality in the labor market. We construct a four-stage model of occupational segregation between two homogeneous, exogenously given, mutually exclusive social groups acting in a two-job labor market. In the first stage each individual chooses one of two specialized educations to become a worker. In the second stage individuals randomly form "friendship" ties with other individuals, with a tendency to form relatively more ties with members of the same social group, what is known in the literature as "*(inbreeding) homophily*", "*inbreeding bias*" or "*assortative matching*".[2] In the third stage workers use their networks of friendship contacts to search for jobs. In the fourth stage workers earn a wage and spend their income on a single consumption good.

We obtain the following results. First, and not surprisingly, we show that with inbreeding homophily within social groups, a complete polarization in terms of occupations across the two groups arises as a stable equilibrium outcome. This result follows from standard arguments on network effects. If a group is completely segregated and specialized in one type of job, then each individual in the group has many more job contacts if she "sticks" to her specialization. Hence, sticking to one specialization ensures good job opportunities to the group members, and these incentives stabilize segregation.

We next extend the basic model allowing for "good" and "bad" jobs, in order to analyze equilibrium wage and unemployment inequality between the two social groups. We show that with large differences in job attraction (=wages), the main outcome of the model is that one social group "fully specializes" in the good job, while the other group "mixes" over the two jobs. In this partial segregation equilibrium, the group that specializes in the good job always has a higher payoff and a lower unemployment rate. Furthermore, with a sufficiently large intra-group homophily, the fully-specializing group also has a higher equilibrium employment rate *and* a higher wage rate than the "mixing" group, thus being twice advantaged. Hence, our model is able to explain typical empirical patterns of gender, race or ethnic labor market inequality. The driving force behind our result is the fact that the group that fully specializes, being homogenous occupationally, is able to create a denser job contact network than the mixing group.

We finally consider whether society benefits from an integration policy, in that labor inequality between the social groups is attenuated. To this aim, we analyze a social plan-

[2]Homophily measures the relative frequency of within-group versus between-group friendships. There exists inbreeding homophily or an inbreeding bias if the group's homophily is higher than what would have been expected if friendships are formed randomly. See Currarini, Jackson and Pin (2008) for formal definitions.

ner's first and second-best policy choices. Surprisingly, segregation is the preferred outcome in the first-best analysis, while a laissez-faire policy leading to segregation shaped by individual incentives is maximizing social welfare in the second-best case. Hence, overall employment is higher under segregation, while laissez-faire inequality remains sufficiently constrained so that segregation is an overall socially optimal policy. Our social welfare analysis points out therefore some relevant policy issues typically ignored in debates concerning anti-segregation legislature.

This paper is mostly related to the segregation framework of Roland Benabou (1993).[3] Benabou introduces a model in which individuals choose between high and low education. The benefits of education, wages, are determined in the global labor market, but the costs are determined by local education externalities. In particular, the costs of high education are considerably more reduced than the costs of low education if many neighbors are highly educated as well, leading to underinvestment in education in the low-education neighborhoods. Benabou shows that these local education externalities lead to segregation and also to inequality at the macro level.[4] Our model is a version of that of Benabou: the link between local externalities and global outcomes is modeled similarly. However, there are a few essential modeling differences leading to markedly different implications. In Benabou (1993) agents choose *different* education levels, either high or low, and thus the marginal productivity and the wage are naturally higher for high-educated workers. Hence, in a segregation equilibrium the highly educated group (or neighborhood, in Benabou's model) has a natural wage advantage. This implies that, under the education externality mechanism, differences in education levels should fully explain the wage gap. As we discuss in more detail in Section 4.2.1, though there is evidence that ability and education differences may explain to a considerable extent the racial wage gap, these differentials cannot fully account for the gender wage gap. Moreover, in Benabou there is no involuntary unemployment and therefore unemployment differences (between races/genders/ethnicities) remain unexplained.

The main difference between this model and Benabou's concerns the results on social welfare. Whereas Benabou suggests that under education externalities *integration* may be the socially optimal policy, we argue here, in contrast, that a social planner would like to *segregate* society. The reason for this different outcome is that the education externalities flow only from high to low education in Benabou's framework—low educated agents

[3]The precursor of many studies on segregation is the seminal work by Schelling (1971), on the emergence of neighborhood racial segregation from tiny differences in the tolerance threshold levels of members of each of two races, regarding the presence of members of the other race.

[4]Furthermore, by introducing the option to drop out of the labor market, Benabou shows that some neighborhoods may turn into ghettos of drop-outs, and this has a dramatic impact on total welfare.

"learn" from high educated agents—whereas externalities are symmetric in our model and thus equally beneficial to both groups. Intuitively, in Benabou (1993) segregation harms the group that has no high-educated agents and this group is better off by enforcing integration. On the other hand, in this paper contact networks are always more effective *for both groups* when there is segregation. Our paper thus shows that a subtle difference in the mechanism of the local externalities can have major implications on optimal social policy.

Significant progress has been lately achieved in modeling labor market phenomena by means of social networks. Recent articles have for instance investigated the effect of social networks on employment, wage inequality, and labor market transitions.[5] This work points out that individual performance on the labor market crucially depends on the position individuals take in the social network structure. However, these studies typically do not focus on the role that networks play in accounting for persistent patterns of occupational segregation and inequality between races, genders or ethnicities.[6] Here, instead of focusing on the network structure, we take a simple reduced form approach, and we emphasize the mechanism relating the role of the job networks in the labor market to occupational segregation and inequality between social groups.

The paper is organized as follows. The next section shortly overviews empirical findings on occupational segregation. We review empirical evidence on the relevance of job contact networks and the extent of social group homophily in Section 2; we set up our model of occupational segregation in Section 4.3; and we discuss key results on the segregation equilibria in Section 4.4. Section 4.5 analyses the social welfare outcome. We summarize and conclude the paper in Section 4.6.

4.2 Empirical background

In this section we present the empirical background that motivates the building blocks of our model. We first discuss evidence on occupational segregation, and the relation to gender and race wage gaps. Next we overview some empirical literature on the role of job contact networks and on homophily.

[5] The seminal paper on the role of networks in the labor markets is Montgomery (1991). Recent papers include Arrow and Borzekowski (2004), Calvó-Armengol and Jackson (2004, 2007), Fontaine (2008), Lavezzi and Meccheri (2005), Bramoullé and Saint-Paul (2006), Ioannides and Soutevent (2006).

[6] Calvó-Armengol and Jackson (2004) find that two groups with two different networks may have different employment rates due to the endogenous decision to drop out of the labor market. However, their finding draws heavily on an example that already assumes a large amount of inequality; in particular, the groups are initially unconnected and the initial employment state of the two groups is unequal.

4.2.1 The extent of occupational segregation

Although labor markets have become more open to traditionally disadvantaged groups, wage differentials by race and gender remain stubbornly persistent. Altonji and Blank (1999) give an overview of the literature on this topic. They note for instance that in 1995 a full-time employed white male earned on average $ 42,742, whereas a full-time employed black male earned on average $ 29,651, thus 30% less, and an employed white female $ 27,583, that is, 35% less. Standard wage regressions are typically able to explain only half of this gap, but more detailed analysis reveals more insights. In particular, several authors have found that the inclusion of individual scores at the Armed Forces Qualifying Test is able to fill the wage gap on race, see the discussion in Altonji and Blank (1999) and the references therein. On the one hand, this suggests that the gap between whites and blacks is created before individuals enter the labor market. On the other hand, the gender wage gap cannot be fully accounted for by pre-market factors, as men and women usually have similar levels of education nowadays.

Much research within social sciences suggests that segregation into separate type of jobs, i.e. occupational segregation, explains a large part of the gender wage gap, as well as part of the race wage gap. A few examples of studies that review and/or present detailed statistics on the occupational segregation[7] and wage inequality patterns by gender, race or ethnicity are Beller (1982), Albelda (1986), King (1992), Padavic and Reskin (2002), Charles and Grusky (2004). All these studies agree that, despite substantial expansion in the labor market participation of women and affirmative action programs aimed at labor integration of racial and ethnic minorities, women typically remain clustered in female-dominated occupations, while blacks and several other races and ethnic groups are over-represented in some occupations and under-represented in others; these occupations are usually of lower 'quality', meaning they are paying less on average, which explains partly the male-female and white-black wage differentials[8].

King (1992) offers for instance detailed evidence that throughout 1940-1988 there was a persistent and remarkable level of occupational segregation by race and sex, such

[7]Some of these papers, e.g. Sørensen (2004), discuss in detail the extent of labor market segregation between social groups, at the *workplace, industry and occupation* levels. Here we shall be concerned with modeling segregation by *occupation* alone (known also as "horizontal segregation"), which appears to be dominant at least relative to segregation by industry. Weeden and Sørensen (2004) convincingly show that occupational segregation in the USA is much stronger than segregation by industries and that if one wishes to focus on one single dimension, "occupation is a good choice, at least relative to industry".

[8]The other prominent side of the 'labor market segregation explaining the wage penalty' story is that women relative to men and, respectively, blacks vis-à-vis whites might experience wage differentials within the same occupation, when located in different workplaces; then we deal with the so-called *vertical segregation* dimension. As stated above, we shall be concerned in this paper only with the occupational dimension, i.e. *horizontal segregation*.

that "approximately two-thirds of men or women would have to change jobs to achieve complete gender integration", with some changes in time for some subgroups. Whereas occupational segregation between white and black women appears to have diminished during the 60's and the 70's, occupational segregation between white and black males or between males and females remained remarkable stable. Several studies by Barbara Reskin and her coauthors, c.f. the discussion and references in Padavic and Reskin (2002), document the extent of occupational segregation by narrow race-sex-ethnic cells and find that segregation by gender remained extremely prevalent and that within occupations segregated by gender, racial and ethnic groups are also aligned along stable segregation paths. Though most of these studies are for the USA, there is also international evidence (particularly from Europe) confirming that, with some variations, similar patterns of segregation hold, e.g. Pettit and Hook (2005).

4.2.2 Job contact networks

There is by now an established set of facts showing the importance of the informal job networks in matching job seekers to vacancies. For instance, on average about 50 percent of the workers obtain jobs through their personal contacts, e.g. Rees (1966), Granovetter (1995), Holzer (1987), Montgomery (1991), Topa (2001); Bewley (1999) enumerates several studies published before the 90's, where the fraction of jobs obtained via friends or relatives ranges between 30 and 60 percent[9]. It is also established that on average 40-50 percent of the employers actively use social networks of their current employees to fill their job openings, e.g. Holzer (1987). Furthermore, employer-employee matches obtained via contacts appear to have some common characteristics. Those who found jobs through personal contacts were on average more satisfied with their job, e.g. Granovetter (1995), and were less likely to quit, e.g. Datcher (1983), Devine and Kiefer (1991), Simon and Warner (1992), Datcher Loury (2006). For a more detailed overview of studies on job information networks, Ioannides and Datcher Loury (2004) is a recent reference.

[9] The difference in the use of informal job networks among professions is also documented. Granovetter (1995) pointed out that although personal ties seem to be relevant in job search-match for all professions, their incidence is higher for blue-collar workers (50 to 65 percent) than for white-collar categories such as accountants or typists (20 to 40 percent). However, for certain other white-collar categories, the use of social connection in job finding is even higher than for blue-collars, e.g. as high as 77 percent for academics.

4.2.3 Intra-group homophily

There is considerable evidence on the existence of the so-called social “homophily”[10], also labeled “assortative matching” or “inbreeding social bias”, that is, there is a higher probability of establishing links among people with similar characteristics. Extensive research shows that people tend to be friends with similar others, see for instance McPherson, et al. (2001) for a review, with characteristics such as race, ethnicity or gender being essential dimensions of homophily. It has also been documented that friendship patterns are more homophilous than would be expected by chance or availability constraints, even after controlling for the unequal distribution of races or sexes through social structure, e.g. Shrum, Cheek and Hunter (1988). There are also studies pointing towards "pure" same race preferences in marrying or dating (e.g. the “mating taboo” in Wong 2003 or the speed dating preferences in Fisman et al. 2006), among very young kids (e.g. Hraba and Grant 1970) or among audiences of television shows (Dates 1980, Lee 2006).

In our "job information network" context, early studies by Rees (1966) and Doeringer and Piore (1971) showed that workers who had been asked for references concerning new hires were in general very likely to refer people "similar" to themselves. While these similar features could be anything, such as ability, education, age, race and so on, the focus here is on groups stratified along exogenous characteristics (i.e. one is born in such a group and cannot alter her group membership) such as those divided along gender, race or ethnicity lines. Indeed, most subsequent evidence on homophily was in the context of such ’exogenously given’ social groups. For instance, Marsden (1987) finds using the U.S. General Social Survey that personal contact networks tend to be highly segregated by race, while other studies such as Brass (1985) or Ibarra (1992), using cross-sectional single firm data, find significant gender segregation in personal networks. Recent evidence is also given by Mayer and Puller (2008) and Currarini, et al. (2008).

Direct evidence of large gender homophily within job contact networks comes from tabulations in Montgomery (1992). Over all occupations in a US sample from the National Longitudinal Study of Youth, 87 percent of the jobs men obtained through contacts were based on information received from other men and 70 percent of the jobs obtained informally by women were as result of information from other women. Montgomery shows that these outcomes hold even when looking at each narrowly defined occupation categories or one-digit industries[11], including traditionally male or female dominated occupations,

[10]The "homophily theory" of friendship was first introduced and popularized by the sociologists Paul F. Lazarsfeld and Robert K. Merton (1954).

[11]Weeden and Sørensen (2004) estimate a two-dimensional model of gender segregation, by industry *and* occupation: they find much stronger segregation across occupations than across industries. 86% of the total association in the data is explained by the segregation along the occupational dimension; this

where job referrals for the minority group members were obtained still with a very strong assortative matching via their own gender group. For example, in male-dominated occupations such as machine operators, 81 percent of the women who found their job through a referral, had a female reference. Such figures are surprisingly large and are likely to be only lower bounds for magnitudes of inbreeding biases within other social groups[12].

Another relevant piece of evidence is the empirical study by Fernandez and Sosa (2005) who use a dataset documenting both the *recruitment* and the *hiring* stages for an entry-level job at a call center of a large US bank. This study also finds that contact networks contribute to the gender skewing of jobs, in addition documenting directly that there is strong evidence of gender homophily in the refereeing process: referees of both genders tend to strongly produce same sex referrals.

Finally, we briefly address the relative importance of homophily within "exogenously given" versus "endogenously created" social groups. As mentioned above, assortative matching takes place along a great variety of dimensions. However, there is empirical literature suggesting that homophily within exogenous groups such as those divided by race, ethnicity, gender, and- to a certain extent- religion, typically outweighs assortative matching within endogenously formed groups such as those stratified by educational, political or economic lines. E.g., Marsden (1988) finds for US strong inbreeding bias in contacts between individuals of the same race or ethnicity and less pronounced homophily by education level. Another study by Tampubolon (2005), using UK data, documents the dynamics of friendship as strongly affected by gender, marital status and age, but not by education, and only marginally by social class. These facts motivate why we focus here on "naturally" arising social groups, such as gender, racial or ethnic ones; nevertheless, as will become clear in the modeling, assuming assortative matching by education, *in addition* to gender, racial or ethnic homophily, does not matter for our conclusions.

4.3 A model of occupational segregation

Based on the stylized facts mentioned in Section 4.2.2, we build a parsimonious theoretical model of social network interaction able to explain stable occupational segregation, and employment and wage gaps, without a need for alternative theories.

Let us consider the following setup. A continuum of individuals with measure 1 is

increases to about 93% once industry segregation is also accounted for. See also footnote 8.

[12]The gender homophily is likely to be smaller than race or ethnic homophily, given frequent close-knit relationships between men and women. This is confirmed for instance by Marsden (1988), who finds strong inbreeding biases in contacts between individuals of the same race or ethnicity, but less pronounced homophily within gender categories.

equally divided into two social groups, Reds (R) and Greens (G). The individuals are ex ante homogeneous apart from their social color. They can work in two occupations, A or B. Each occupation requires a corresponding thorough specialized education (career track), such that a worker cannot work in it unless she followed that education track. We assume that it is too costly for individuals to follow both educational tracks. Hence, individuals have to choose their education track before they enter the labor market.[13]

Consider now the following order of events:

1. Individuals choose one education in order to specialize either in occupation A or in occupation B;
2. Individuals randomly establish "friendship" relationships, thus forming a network of contacts;
3. Individuals participate in the labor market. Individual i obtains a job with probability s^i.
4. Individuals produce a single good for their firms and earn a wage w^i. They obtain utility from consuming goods that they buy with their wage.

We proceed with an elaboration of these steps.

4.3.1 Education strategy and equilibrium concept

The choice of education in the first stage involves strategic behavior. Workers choose the education that maximizes their expected payoff given the choices of other workers, and we therefore look for a Nash equilibrium in this stage. This can be formalized as follows.

Denote by μ_R and μ_G the fractions of Reds and respectively Greens that choose education A. It follows that fraction $1 - \mu_X$ of group $X \in \{R, G\}$ chooses education B. The payoffs will depend on these strategies: the payoff of a worker of group X that chooses education A is given by $\Pi_A^X(\mu_R, \mu_G)$, and mutatis mutandis, $\Pi_B^X(\mu_R, \mu_G)$. Define $\Delta\Pi^X \equiv \Pi_A^X - \Pi_B^X$. The functional form of the payoffs is made more specific later, in subsection 4.3.4.

In a Nash equilibrium each worker chooses the education that gives her the highest payoff, given the education choices of all other workers. Since workers of the same social group are homogenous, a Nash equilibrium implies that if some worker in a group chooses education A (B), then no other worker in the same group should prefer education B (A). This implies that a pair (μ_R, μ_G) is an *equilibrium* if and only if, for $X \in \{R, G\}$, the

[13] For example, graduating high school students may face the choice of pursuing a medical career or a career in technology. Both choices require several years of expensive specialized training, and this makes it unfeasible to follow both career tracks.

following hold:[14]

$$\Delta\Pi^X(\mu_R, \mu_G) \leq 0 \text{ if } \mu_X = 0 \tag{4.1}$$

$$\Delta\Pi^X(\mu_R, \mu_G) = 0 \text{ if } 0 < \mu_X < 1 \tag{4.2}$$

$$\Delta\Pi^X(\mu_R, \mu_G) \geq 0 \text{ if } \mu_X = 1. \tag{4.3}$$

To strengthen the equilibrium concept, we restrict ourselves to *stable equilibria.* We use a simple stability concept based on a standard myopic adjustment process of strategies, which takes place before the education decision is made. That is, we think of the equilibrium as the outcome of an adjustment process. In this process, individuals repeatedly announce their preferred education choice, and more and more workers revise their education choice if it is profitable to do so, given the choice of the other workers.[15] Concretely, we consider stationary points of a dynamic system guided by the differential equation $\dot{\mu}_X = k\Delta\Pi^X(\mu_R, \mu_G)$. This implies that $\mu \equiv (\mu_R, \mu_G)$ is a stable equilibrium if it is an equilibrium and (i) for $X \in \{R, G\}$: $\partial\Delta\Pi^X/\partial\mu_X < 0$ if $\Delta\Pi^X = 0$; (ii) $\det(D\Delta\Pi(\mu)) > 0$ if $\Delta\Pi^R = 0$ and $\Delta\Pi^G = 0$, where $D\Delta\Pi(\mu)$ is the Jacobian of $(\Delta\Pi^R, \Delta\Pi^G)$ with respect to μ.

4.3.2 Network formation

In the second stage the workers form a network of contacts. We assume this network to be random, but with social color homophily. That is, we assume that the probability for two workers to create a tie is $p \geq 0$ when the two workers are from different social groups and follow different education tracks; however, when the two workers are from the same social group, the probability of creating a tie increases with $\lambda > 0$. Similarly, if two workers choose the same education, then the probability of creating a tie increases with $\kappa \geq 0$. Hence, we allow for assortative matching by education, in addition to the one by social color. We do not impose any further restrictions on these parameters, other than securing $p + \lambda + \kappa \leq 1$. This leads to the tie formation probabilities from Table 4.1. We shall refer to two workers that create a tie as "friends"

[14]The question whether the equilibrium is in pure or mixed strategies is not relevant, because the player set is a measure of identical infinitesimal individuals (except for group membership). Our equilibrium could be interpreted as a Nash equilibrium in pure strategies; then μ_X is the measure of players in group X choosing pure strategy A. The equilibrium could also be interpreted as a symmetric Nash equilibrium in mixed strategies; in that case the common strategy of all players in group X is to play A with probability μ_X. A hybrid interpretation is also possible.

[15]One could think of such a process as the discussions students have before the end of the high school about their preferred career. An alternative with a longer horizon is an overlapping generations model, in which the education choice of each new generation partly depends on the choice of the previous generation.

Table 4.1: The probability of a tie between two individuals, depending on the group membership and education choice.

		Education	
		same	different
Social group	same	$p+\kappa+\lambda$	$p+\lambda$
	different	$p+\kappa$	p

We assume the probability that an individual i forms a tie with individual j to be exogenously given and constant. In practice, establishing a friendship between two individuals typically involves rational decision making. It is therefore plausible that individuals try to optimize their job contact network in order to maximize their chances on the labor market.[16] In particular, individuals from the disadvantaged social groups should have an incentive to form ties with individuals from the advantaged group. While this argument is probably true, we do not incorporate this aspect of network formation in our model. The harsh reality is that strategic network formation does not appear to dampen the inbreeding bias in social networks significantly; in Section 4.2.2 we provided an abundance of evidence that strong homophily exists even within groups that have strong labor market incentives *not* to preserve such homophily in forming their ties. The reason could be that the payoff of forming a tie is mainly determined by various social and cultural factors, and only for a smaller part by benefits from the potential transmission of valuable job information.[17] On top of that, studies such as, for instance, Granovetter (2002), also note that many people would feel exploited if they find out that someone befriends them for the selfish reason of obtaining job information. These elements might hinder the role of labor market incentives when forming ties. Hence, while we do not doubt that incentives do play a role when forming ties, we believe these incentives are not sufficient to undo the effects of the social color homophily. We therefore assume network formation exogenous in this paper.

4.3.3 Job matching and social networks

The third stage we envision for this model is that of a dynamic labor process, in which information on vacancies is propagated through the social network, as in, e.g., Calvó-

[16] See Calvó-Armengol (2004) for a model of strategic network formation in the labor market.

[17] Currarini, et al.(2008) discuss a model of network formation in which individuals form preferences on the number and mix of same-group and other-group friends. In this model inbreeding homophily arises endogenously.

Armengol and Jackson (2004), Calvó-Armengol and Zenou (2005), Ioannides and Soetevent (2006) or Bramoullé and Saint-Paul (2006). Workers who randomly lose their job are initially unemployed because it takes time to find information on new jobs. The unemployed worker receives such information either directly, through formal search, or indirectly, through employed friends who receive the information and pass it on to her (in the particular case where all her friends are unemployed, only the formal search method works). As the specific details of such a process are not important for our purposes, we do not consider these dynamic models explicitly, but take a "reduced form" approach.

In particular, we assume that unemployed workers have a higher propensity to receive job information when they have more friends with the *same job background*, that is, with the *same choice of education*. On the one hand, this assumption is based on the result of Ioannides and Soetevent (2006) that in a random network setting the individuals with more friends have a lower unemployment rate.[18] On the other hand, this assumption is based on the conjecture that workers are more likely to receive information about jobs in their own occupation. For example, when a vacancy is opened in a team, the other team members are the first to know this information, and are also the ones that have the highest incentives to spread this information around.

Formally, denote the probability that individual i becomes employed by $s^i = s(x_i)$, where x_i is the measure of friends of i with the same education as i has. We thus assume that $s(x)$ is differentiable, $0 < s(0) < 1$ (there is non-zero amount of direct job search) and $s'(x) > 0$ for all $x > 0$ (the probability of being employed increases in the number of friends with the same education).

It is instructive to show how s^i depends on the education choices of i and the choices of all other workers. Remember that μ_R and μ_G are the fractions of Reds and respectively Greens that choose education A. Given the tie formation probabilities from Table 4.1 and some algebra, the employment rate s_A^X of A-workers in group $X \in \{R, G\}$ will be given by:

$$s_A^X(\mu_R, \mu_G) = s\left((p+\kappa)\bar{\mu} + \lambda\mu_X/2\right) \tag{4.4}$$

and likewise, the employment rate s_B^X of B-workers in group X will be

$$s_B^X(\mu_R, \mu_G) = s\left((p+\kappa)(1-\bar{\mu}) + \lambda(1-\mu_X)/2\right) \tag{4.5}$$

[18] This result is nontrivial, as the unemployed friends of employed individuals tend to compete with each other for job information. Thus, if a friend of a jobseeker has more friends, the probability that this friend passes information to the jobseeker decreases. In fact, in a setting in which everyone has the same number of friends, Calvó-Armengol and Zenou (2005) show that the unemployment rate is non-monotonic in the (common) number of friends.

where $\bar{\mu} \equiv (\mu_R + \mu_G)/2$.

Note that $s_A^X > s_A^Y$ and $s_B^X < s_B^Y$ for $X, Y \in \{R, G\}$, $X \neq Y$, if and only if $\mu_X > \mu_Y$ and $\lambda > 0$. We will see in Section 4.4.1 that the ranking of the employment rates is crucial, as it creates a group-specific network effect. That is, keeping this ordering, if only employment matters (jobs are equally attractive), then individuals have an incentive to choose the same education as other individuals in *their* social group. Importantly, it is straightforward to see that this ordering of the employment rates depends on λ, but it does not depend on κ. Therefore, only the homophily among members of the same social group- and not the eventual assortative matching by education- is relevant to our results.

4.3.4 Wages, consumption and payoffs

The eventual payoff of the workers depends on the wage they receive, the goods they buy with that wage, and the utility they derive from consumption. Without loss of generality we assume that an unemployed worker receives zero wage. However, the wages of employed workers are not exogenously given, but they are determined by supply and demand.

When firms offer wages, they take into account that there are labor market frictions and that it is impossible to employ all workers simultaneously. Thus what matters is the effective supply of labor as determined by the labor market process in stage 3. Let L_A be the total measure of employed A-workers and L_B be the total measure of employed B-workers. Hence,

$$L_A(\mu_R, \mu_G) = \mu_R s_A^R(\mu_R, \mu_G)/2 + \mu_G s_A^G(\mu_R, \mu_G)/2 \tag{4.6}$$

and

$$L_B(\mu_R, \mu_G) = (1 - \mu_R) s_B^R(\mu_R, \mu_G)/2 + (1 - \mu_G) s_B^G(\mu_R, \mu_G)/2. \tag{4.7}$$

Given (4.4) and (4.5) from above, it is easy to check that L_A is increasing with μ_R and μ_G, whereas L_B is decreasing with μ_R, μ_G.

As in Benabou (1993), consumption, prices, utility, the demand for labor and the implied wages are determined in a 1-good, 2-factor general equilibrium model. All individuals have the same utility function $U : R_+ \to R$, which is strictly increasing and strictly concave with $U(0) = 0$. The single consumer good sells at unit price, such that consumption of this good equals wage and indirect utility is given by $U_i = U(w_i)$.

Firms put A-workers and B-workers together to produce the single good at constant returns to scale. Wages are then determined by the production function $F(L_A, L_B)$. As usually, we assume that F is strictly increasing and strictly concave in L_A and L_B and $\partial^2 F/\partial L_A \partial L_B > 0$. Writing the wage as function of education choices and using (4.6) and

(4.7), the wages of A-workers and B-workers, w_A and w_B, are given by

$$w_A(\mu_R, \mu_G) = \frac{\partial F}{\partial L_A}\left(L_A(\mu_R, \mu_G), L_B(\mu_R, \mu_G)\right),$$

and

$$w_B(\mu_R, \mu_G) = \frac{\partial F}{\partial L_B}\left(L_A(\mu_R, \mu_G), L_B(\mu_R, \mu_G)\right).$$

It is easy to check that w_A is strictly decreasing with μ_R and μ_G, and mutatis mutandis, w_B.

We can now define the payoff of a worker as her expected utility at the time of decision-making. The payoff function of an A-educated worker from social group $X \in \{R, G\}$ is thus

$$\Pi_A^X(\mu_R, \mu_G) = s_A^X(\mu_R, \mu_G)U(w_A(\mu_R, \mu_G)). \tag{4.8}$$

Similarly,

$$\Pi_B^X(\mu_R, \mu_G) = s_B^X(\mu_R, \mu_G)U(w_B(\mu_R, \mu_G)). \tag{4.9}$$

If we do not impose further restrictions, then there could be multiple equilibria, most of them uninteresting. To ensure a unique equilibrium in our model (actually: two symmetric equilibria), we make the following two assumptions.

Assumption 1 *For the wage functions w_A and w_B*

$$\lim_{x \downarrow 0} U(w_A(x,x)) = \lim_{x \downarrow 0} U(w_B(1-x, 1-x)) = \infty.$$

Assumption 2 *For $X \in \{R, G\}$, and for all $\mu_R, \mu_G \in [0, 1]$*

$$\left|\frac{\partial s_A^X / s_A^X}{\partial \mu_X / \mu_X}\right| < \left|\frac{\partial U/U}{\partial w_A / w_A}\right| \left|\frac{\partial w_A / w_A}{\partial \mu_X / \mu_X}\right|$$

and

$$\left|\frac{\partial s_B^X / s_B^X}{\partial \mu_X / \mu_X}\right| < \left|\frac{\partial U/U}{\partial w_B / w_B}\right| \left|\frac{\partial w_B / w_B}{\partial \mu_X / \mu_X}\right|.$$

Assumptions 1 and 2 guarantee the uniqueness of our results. Assumption 1 implies that the wage for scarce labor is so high that at least some workers always find it attractive to choose education A or respectively B; everyone going for one of the two educations cannot be an equilibrium. In Assumption 2 we assume that the education choice of an individual has a smaller marginal effect on the employment probability within a group than on the wages and overall utility. Note that the assumption implies that for $X \in \{R, G\}$

$$\frac{\partial \Pi_A^X}{\partial \mu_X} < 0 < \frac{\partial \Pi_B^X}{\partial \mu_X},$$

and it is this feature that guarantees the uniqueness of our results. The assumption is not restrictive as long as there is sufficient direct job search, because the employment probability of each individual in our model is bounded between $s(0) > 0$ and 1, with $s(0)$ capturing the employment probability in the absence of any ties and thus induced only by the exogenously given direct job finding rate. Hence, a higher $s(0)$ implies less of an impact of the network effect on the employment rate.

It should be noted that we make these assumptions above only in order to focus our analysis on segregation outcomes, for the sake of clarity and brevity. These assumptions are not necessary. For instance, in the calibration exercise of Section 4.5.2, Assumption 2 is violated, but there are still (two) unique equilibria.

4.4 Equilibrium results

We now present the equilibrium analysis of our model. The formal proofs of all subsequent propositions are relegated to the Appendix. Without loss of generality we assume throughout the section that $w_A(1,0) \geq w_B(1,0)$, thus that the A-occupation is weakly more attractive than the B-occupation when effective labor supply is equal. We call A the "good" job, and B the "bad" job.

4.4.1 Occupational segregation

We are in particular interested in those equilibria in which there is segregation. We define *complete segregation* if $\mu_R = 0$ and $\mu_G = 1$, or, vice versa, $\mu_R = 1$ and $\mu_G = 0$. On the other hand, we say that there is *partial segregation* if for $X \in \{R, G\}$ and $Y \in \{R, G\}$, $Y \neq X$: $\mu_X = 0$ but $\mu_Y < 1$, or, vice versa, $\mu_X = 1$ but $\mu_Y > 0$.

Our first result is that segregation, either complete or partial, is the only stable outcome:

Proposition 3 *Suppose Assumptions 1 and 2 hold. Define* $s_H \equiv s((p + \kappa + \lambda)/2)$ *and* $s_L \equiv s((p + \kappa)/2)$.

(i) If

$$1 \leq \frac{U(w_A(1,0))}{U(w_B(1,0))} \leq \frac{s_H}{s_L}, \tag{4.10}$$

then there are exactly two stable equilibria, both with complete segregation.

(ii) If

$$\frac{U(w_A(1,0))}{U(w_B(1,0))} > \frac{s_H}{s_L}, \tag{4.11}$$

then there are exactly two stable equilibria, both with partial segregation, in which either $\mu_R = 1$ or $\mu_G = 1$.

We first note that a non-segregation equilibrium cannot exist, even in the case of a tiny amount of homophily (λ very small). The intuition is that homophily in the social network among members of the same social group creates a group-dependent network effect. Thus, if slightly more Red workers choose A than Greens do, then the value of an A-education is higher for the Reds than for the Greens, while the value of a B-education is lower in the Reds' group. Positive feedback then ensures that the initially small differences in education choices between the two groups widen and widen, until at least one group segregates completely into one type of education.

Second, if the wage differential between the two jobs (for equal numbers of A-educated and B-educated workers) is not "too large" vis-à-vis the social network effect (condition 4.10), complete segregation is the only stable equilibrium outcome, given a positive inbreeding bias in the social group. Thus one social group specializes in one occupation, and the other group in the other occupation. On the other hand, the proposition makes clear that complete segregation cannot be sustained if the wage differential is "too large" vis-à-vis the social network effect (condition 4.11). Starting from complete segregation, a large wage differential gives incentives to the group specialized in B-jobs to switch to A-jobs.

Interestingly, the "unsustainable" complete segregation equilibrium is then replaced by a partial equilibrium in which one group specializes in the "good" job A, while the other group has both A and B-workers. Partial segregation in which one group, say the Greens, fully specializes in the "bad" job B is unsustainable, as that would lead to an oversupply of B-workers and an even larger wage differential. This would provide the Red B-workers with strong incentives to switch en masse to the A-occupation.

4.4.2 Inequality

The discussion so far ignored eventual equilibrium differentials in wages and unemployment between the two types of jobs. We now tackle that case. We continue to assume that $w_A(1,0) \geq w_B(1,0)$ and, in light of the results of Proposition 3, we focus without loss of generality on the equilibrium in which $\mu_R = 1$. Thus, the Reds specialize in the "good" job A, while the "bad" job B is only performed by Green workers.

We first consider the case in which wage differentials are small enough so that complete segregation is an equilibrium ($\mu_R = 1$ and $\mu_G = 0$). In this case the implications are straightforward. Since both groups specialize in equal amounts, the network effects are equally strong, and the employment rates are equal. Given that employment rates are equal, the effective labor supply is also equal, and therefore the wage of the "good" job is weakly higher. We thus have the following result:

Proposition 4 *Suppose Assumptions 1 and 2 hold. Define* $s_H \equiv s((p+\kappa+\lambda)/2)$ *and* $s_L \equiv s((p+\kappa)/2)$ *and suppose that* $1 \leq \frac{w_A(1,0)}{w_B(1,0)} \leq \frac{s_H}{s_L}$*. Suppose* $(\mu_R, \mu_G) = (1,0)$ *is a stable equilibrium. In that equilibrium*

$$w_A \geq w_B,$$

$$s_A^R = s_B^G > s_B^R = s_A^G,$$

and

$$\Pi_A^R \geq \Pi_B^G \geq \Pi_A^G \geq \Pi_B^R. \tag{4.12}$$

This result is not very surprising, hence we turn next to the analysis of the more interesting case in which wage differentials are large. In that case there is a partial equilibrium in which $(\mu_R, \mu_G) = (1, \mu^*)$ where $\mu^* \in (0,1)$. First note that according to (4.2) this implies the following condition:

$$\Pi_A^G(1,\mu^*) = \Pi_B^G(1,\mu^*),$$

or equivalently

$$s_A^G(1,\mu^*)U(w_A(1,\mu^*)) = s_B^G(1,\mu^*)U(w_B(1,\mu^*)).$$

Thus, whereas workers in group R prefer the A-job, the workers in group G make an individual trade-off: lower wages should be exactly compensated by higher employment probabilities and vice versa.

We are particularly interested in whether this individual trade-off between unemployment and wages translates into a similar trade-off at the 'macro-level', in which an inter-group wage gap is compensated by a reversed employment gap. We have the following proposition.

Proposition 5 *Suppose Assumptions 1 and 2 hold. Define* $s_H \equiv s((p+\kappa+\lambda)/2)$ *and* $s_L \equiv s((p+\kappa)/2)$ *and suppose that* $\frac{w_A(1,0)}{w_B(1,0)} > \frac{s_H}{s_L}$*. Define* $\hat{\mu} \in (0,1)$*, such that*

$$w_A(1,\hat{\mu}) = w_B(1,\hat{\mu}), \tag{4.13}$$

and let $(\mu_R, \mu_G) = (1, \mu^*)$ *be a stable equilibrium. In that equilibrium*

$$\Pi_A^X > \Pi_B^Y = \Pi_A^Y > \Pi_B^X. \tag{4.14}$$

Moreover,

(i) if $\hat{\mu} < \frac{\lambda}{2(p+\kappa+\lambda)}$, *then*

$$s_A^R > s_B^G > s_A^G > s_B^R,$$

and

$$w_A(1, \mu^*) > w_B(1, \mu^*);$$

(ii) if $\hat{\mu} > \frac{\lambda}{2(p+\kappa+\lambda)}$, *then*

$$s_A^R > s_A^G > s_B^G > s_B^R,$$

and

$$w_B(1, \mu^*) > w_A(1, \mu^*).$$

The main implication of this proposition is that an inter-group wage gap is *not* compensated by a reversed employment gap. On the contrary, it is possible that the group specializing in the good job, here the Reds, both earns a higher wage *and* has higher employment probabilities than the Greens group. This is especially clear when the group homophily bias λ is large relative to p and κ (in fact $p + \kappa$) and there is a big difference in attractiveness between the good and the bad jobs (case (i) above).

This result can be understood by the following observation: the workers in the 'specializing' group R have a higher employment probability than *all* workers in group G. This is always the case, regardless of whether the individual in G is an A or a B worker, and whether $s_B^G > s_A^G$ or not. As all members of group R choose the same occupation, the Reds remain a strong homogenous social group. Network formation with homophily then implies that they are able to create a lot of ties, and hence, that they benefit most from their social network. On the other hand, the Greens are dispersed between two occupations. This weakens their social network and this decreases their chances on the labor market, both for A and B-workers in group G.

Whether the wage differential between the workers in the two groups is positive or negative depends on the relative size of λ relative to $p + \kappa$, in the term $\frac{\lambda}{2(p+\kappa+\lambda)}$ from the inequality conditions in Proposition 5. This can be roughly assessed in light of the empirical evidence on homophily discussed earlier in this paper. First, as seen from the stylized facts from Section 4.2.2, the assortative matching by education, κ, is typically found to be lower relative to racial, ethnical or gender homophily. The second interesting situation is a scenario where the probability of making contacts in general, p, were already extremely high relative to the intra-group homophily bias. However, given the surprisingly large size of intra-group inbreeding biases in personal networks of contacts found empirically, this is also unlikely. Hence, the likelihood is very high that in practice λ would dominate the other parameters in the cutoff term $\frac{\lambda}{2(p+\kappa+\lambda)}$.

Let us sum up the implications of this last proposition. The fully specializing group is always better off in terms of unemployment rate and payoff, independent of either relative or absolute sizes of λ, p and κ (as long as $\lambda > 0$), as shown in Proposition 5. Furthermore, given the observed patterns of social networks discussed in Section 4.2.2, the condition of λ dominant relative to p and κ is likely to be met. This ensures that the group fully specializing in the good job always has a higher wage in the equilibrium than the group mixing over the two jobs, as proved in Proposition 5. Note that this partial segregation equilibrium is in remarkable agreement with observed occupational, wage and unemployment disparities in the labor market between, for instance, males-females or blacks-whites. This suggests that our simple model offers a plausible explanation for major empirical patterns of labor market inequality.

4.5 Social welfare

4.5.1 First best social optimum

In the previous section we observed that individual incentives lead to occupational segregation and wage and unemployment inequality. This suggests that a policy targeting integration may reduce inequality as well, and in fact may just be socially beneficial. This is an argument often used for instance by proponents of positive discrimination. We set out here to analyze the implications of our model from a social planner's point of view.

Consider a utilitarian social welfare function:

$$W(\mu_R, \mu_G) = \mu_R \Pi_A^R/2 + (1-\mu_R)\Pi_B^R/2 + \mu_G \Pi_A^G/2 + (1-\mu_G)\Pi_B^G/2, \tag{4.15}$$

where $\Pi_A^X \equiv \Pi_A^X(\mu_R, \mu_G)$ and $\Pi_B^X \equiv \Pi_B^X(\mu_R, \mu_G)$ are given by equations (4.8) and (4.9). Since unemployed workers obtain zero utility, we can also write the welfare function as

$$W(\mu_R, \mu_G) = L_A U\left(\frac{\partial F}{\partial L_A}(L_A, L_B)\right) + L_B U\left(\frac{\partial F}{\partial L_B}(L_A, L_B)\right), \tag{4.16}$$

where $L_A \equiv L_A(\mu_R, \mu_G)$ and $L_B \equiv L_B(\mu_R, \mu_G)$ were introduced by (4.6) and (4.7). The formulation in (4.16) is useful, because it shows that what matters for social welfare is the effect of a policy on the society's effective labor supply.

We consider a first-best social optimum, that is, the social planner is able to fully manage $\mu_R \in [0,1]$ and $\mu_G \in [0,1]$ and therefore the social optimum $\mu^S = (\mu_R^S, \mu_G^S)$ is defined as

$$\mu^S = \text{argmax}_{\mu_R \in [0,1], \mu_G \in [0,1]} W(\mu_R, \mu_G).$$

We obtain the following result:

Proposition 6 *If for all $x \in [0, (p + \kappa + \lambda)/2]$:*

$$s''(x) > -\frac{4}{\lambda}\, s'(x) \tag{4.17}$$

then any social optima involves complete or partial segregation.

Thus a segregation policy is socially preferred, as long as $s(x)$, the employment probability of having x friends with the same education, is "not too concave". This proposition can be intuitively understood as follows. Suppose that there is no segregation, and $0 < \mu_G < \mu_R < 1$. In that case the Reds obtain a higher employment probability in an A-occupation, $s_A^R > s_B^R$, whereas the Greens have a higher employment rate as B-workers, $s_B^G > s_A^G$. Now consider the effect on segregation, wages and employment when a social planner forces a Red individual initially choosing a B-occupation and respectively, a Green individual initially choosing an A-occupation, into *switching their occupation choice*. In that case μ_R slightly increases, whereas μ_G slightly decreases. The result of this event is, first, that segregation increases; the gap between μ_R and μ_G becomes larger. Second, the total fraction of individuals that choose occupation A, $\mu_R + \mu_G$, does not change. So the ratio of A-workers versus B-workers does not change much, and therefore the ratio of wages does not change much either. Thus the effect on wage inequality is only marginal. Third, by switching occupations, the Red worker can now benefit from a denser network, and have an employment probability s_A^R instead of s_B^R. The same is true for the Green worker switching from B to A. Thus, the combined payoff of the two workers increases, as they are both more likely to become employed. We also need to consider the externality on the employment rates of the workers not involved in the occupation switch. In particular, the switch of occupations increases the network effects of the other Red A-workers and Green B-workers, whereas it decreases the network effects of Red B-workers and Green A-workers. The restriction on the concavity of $s(x)$ ensures that the switch of occupations puts on average a positive externality on the employment probabilities of other workers. We conclude that the switch of occupations of the two workers hardly affects wage inequality, while it increases the labor supply of both A and B. Therefore, social welfare increases.

The general message of this result is that an integration policy might have detrimental effects on employment, effects that are usually overlooked by strong advocates of positive discrimination. Under our model's assumptions, integration might weaken the employment chances of individuals, because the network effects are weaker in mixed networks. In the case of complete segregation, individuals are surrounded by similar individuals during their education. Therefore, it is easier for them to make many friends they can rely on when searching on the job market. Consequently, employment probabilities are high. On the other hand, if educations are mixed, then individuals have more difficulties in creating useful job contacts, and therefore their employment probabilities are lower.

It is worth to point out that the result that integration weakens network effects and decreases labor market opportunities has empirical support in related literature on segregation. For example, Currarini, et al. (2008) find clear evidence that larger (racial) minorities create more friendships, and Marsden (1987) finds a similar pattern in his network of advice. Therefore, it is more beneficial for a worker to choose an education in which she is only surrounded by similar others, instead of an education in which racial groups are mixed, let alone one in which she is a small minority. In a different but related context, Alesina and La Ferrara (2000, 2002) find that participation in social activities is lower in racially mixed communities and so is the level of trust. These and our results suggest that possible negative impacts of integration on social network effects should also be taken into account.

Our outcome on the first-best social optimum hinges for a large part on the fact that the social planner is able to increase employment by increasing segregation, while still controlling wage inequality. In reality however, a social planner may not have this amount of control. Perhaps a more feasible policy is a policy in which the social planner enforces and stabilizes integration, but where the exact allocation of workers to occupations is determined by individual incentives. In the case of segregation there would be a potentially large inequality in payoffs between the social groups, whereas in the case of integration there may be complete payoff equality, but employment may be lower. This suggests a second-best analysis of social welfare, in which there is a potential trade-off of segregation between network benefits and inequality. Such an analysis is unfeasible without further specification of the parameters, hence we will perform that analysis subsequent to calibrating the model for suitable parameters and functional forms.

4.5.2 Second best social optimum

Numerical simulation

As often done in such frameworks, e.g. Fontaine (2008), we calibrate the parameters, in order to perform a small numerical simulation of our model. The purpose of this simulation is to get a better feeling on the mechanisms of the model, the restrictiveness of our assumptions, and the magnitude of the wage gap that can be generated. The simulation also allows us to get some insights about a second-best welfare policy.

We first specify functional forms for $s(x)$, the employment probability as function of the number of friends with the same education, $F(L_A, L_B)$, the production function and thus the derived wage functions, and $U(x)$, the utility function. Regarding the employment probability, we consider a function that follows from a dynamic labor process, in which employed individuals become unemployed at rate 1, and in which unemployed individuals become employed at rate $c_0 + c_1 x$, where c_0 is the rate at which unemployed workers directly obtain information on job vacancies, and c_1 measures the strength of having friends. This leads to the following employment function:

$$s(x) = \frac{c_0 + c_1 x}{1 + c_0 + c_1 x}.$$

Since we have defined $s_0 = s(0)$ as the employment probability when only direct search is used, it follows that $s_0 = c_0/(1 + c_0)$.

For the production function we assume the commonly used Cobb-Douglas function with constant returns to scale,

$$F(L_A, L_B) = \theta L_A^{\alpha} L_B^{1-\alpha}.$$

For the utility function we consider a function with constant absolute risk aversion, where ρ is the coefficient of absolute risk aversion. That is

$$U(x) = 1 - e^{-\rho x}.$$

We calibrate the parameters $s_0, c_1(p + \kappa), c_1\lambda, p$ and θ, leaving α as a free parameter. First, we calibrate s_0, $c_1(p + \kappa)$, and $c_1\lambda$ from three equations that are motivated by the empirical evidence given in Section 2 and 3. This parameterization is sufficient to perform the simulation, and it is thus not necessary to separately specify c_1, p, κ and λ. The first equation is obtained by imposing the restriction that about 50% of the workers find their job through friends, as suggested in Section 2. This restriction implies that the direct job arrival rate c_0 should equal the indirect job arrival rate through friends $c_1 x$. The indirect job arrival rate differs, depending on the choices of the individuals, but if we focus on the case complete segregation, in which $\mu_R = 1$ and $\mu_G = 0$, then we can impose the following restriction:

$$c_0 = c_1(p + \kappa + \lambda)/2.$$

Next, we calibrate the amount of inbreeding homophily in the social group. This amount typically differs depending on the group defining characteristic. For example, analyzing data on Facebook participants at Texas A&M, Mayer and Puller (2008) find that two students living in the same dorm are 13 more likely to be friends than two random students, two black students 17 more likely, but two Asian students 5 times more likely and two Hispanic students twice as likely to be friends. In light of this evidence, we chose to keep the amount of inbreeding homophily in the simulation modest, imposing $\lambda = 3(p + \kappa)$.

We next impose that the employment rate is 95% in case of complete segregation. Given the above, we solve

$$\frac{2c_0}{1 + 2c_0} = 0.95,$$

and this implies that

$$s_0 = \frac{c_0}{1 + c_0} = \frac{19}{21} \approx .9048.$$

and further that $c_1(p + \kappa) = 4.75$ and $c_1\lambda = 14.25$.

Let us consider now the productivity parameter θ and the coefficient of absolute risk aversion ρ. The coefficient of absolute risk aversion has been estimated between 6.6×10^{-5} and 3.1×10^{-4} (Gertner 1993, Metrick 1995, Cohen and Einav 2007). We set the risk aversion at 1.0×10^{-4}, which means a coefficient of relative risk aversion of 4 at a wealth level of \$ 40,000, or indifference at participating in a lottery of getting \$ 100.00 or losing \$ 99.01 with equal probability.

The productivity parameter, θ, is chosen such that average income equal \$ 40,000 in the case of complete segregation, $(\mu_R, \mu_G) = (1, 0)$, and $\alpha = .5$.[19] Since in that situation $w_A(1, 0) = w_B(1, 0) = \theta/2$, we have $\theta = 80,000$.

We can now look at the dependence of payoffs, wages and employment on α with s_0, $c_1(p+\kappa)$, $c_1\lambda$, ρ and θ as summarized in Table 4.2, and in which μ_R and μ_G are determined by equilibrium conditions (4.1)-(4.3). Given the result of Proposition 3 that there is either a complete equilibrium or a partial equilibrium, in which one group specializes in the good job, we concentrate our attention to the parameter space in which $\alpha \in [1/2, 1)$, $\mu_R = 1$

[19]GDP per capita was \$ 44,190 in the U.S. in 2006 according to figures from the IMF.

Table 4.2: Chosen parameter values in the simulation and the sensitivity with respect to $\hat{\alpha}$ and the maximum wage gap.

Parameter	Value	Elasticity of $\hat{\alpha}$ $\hat{\alpha} = .5904$	Elasticity of wage gap $G(1,0) = .306$
s_0	.9048	-1.71	-9.47
$c_1(p+\kappa)$	4.75	-.04	-.23
$c_1\lambda$	14.25	.08	.46
ρ	1.0×10^{-4}	.38	2.09
θ	80,000	.38	2.09

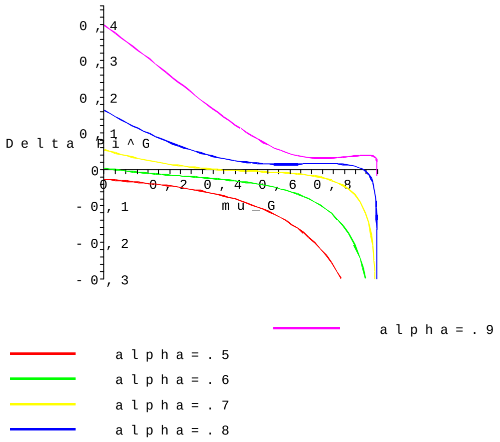

Figure 4.1: $\Delta\Pi^G(1, \mu_G)$ as a function of μ_G for different values of α.

and $\mu_G \in [0, 1)$. Thus occupation A is "good", and group R specializes in A.

We first show a plot of $\Delta\Pi^G(1, \mu_G)$ as a function of μ_G for different values of α. This function illustrates the payoff evaluation that a Green individual makes when deciding on its occupation. If $\Delta\Pi^G(1, \mu_G) > (<)0$, then the Green individual prefers A (B) if she beliefs that all Reds choose A and fraction μ_G of Greens choose A. Clearly, in an equilibrium it should hold that either $\Delta\Pi^G(1, 0) < 0$ or $\Delta\Pi^G(1, \mu_G) = 0$.

The plot is displayed in Figure **??**. This plot nicely illustrates the workings of the model. First, note that for $\alpha = .5$, $\Delta\Pi^G(1, \mu_G)$ is clearly negative, so given that the Reds choose A, the Greens prefer B and complete segregation is an equilibrium. However, $\Delta\Pi^G(1, \mu_G)$ increases with α, such that for $\alpha > .5904 \equiv \hat{\alpha}$, we have that $\Delta\Pi^G(1, 0) > 0$ and complete segregation is not an equilibrium anymore. In that case, there is a unique partial equilibrium.[20]

[20] $\Delta\Pi^G(1, \mu_G)$ is not monotonically decreasing for very large α, which implies that Assumption 2 is

If $\alpha < .5904$ we have complete segregation as an equilibrium. In that case Proposition 4 gives us the employment rates and wages. Employment rates are given by:

$$s_A^R = s_B^G = .95 \text{ and } s_B^R = s_A^G = .9223.$$

Wages have a particular simple form in the case of complete segregation, being $w_A(1,0) = \theta\alpha$ and $w_B(1,0) = \theta(1-\alpha)$. Therefore, if we define the wage gap as $G(\mu_R, \mu_G) = 1 - w_B(\mu_R, \mu_G)/w_A(\mu_R, \mu_G)$, then the wage gap under complete segregation is $G(1,0) = 2 - 1/\alpha$. Note that at $\alpha = \hat{\alpha} = .5904$, we have

$$w_A(1,0) = 47,233 \text{ and } w_B(1,0) = 32,767$$

and the wage gap is thus $G(1,0) = .306$. Hence, a small employment gap of .9223 versus .95 is only compensated by a wage gap of 30 %!

It is worth elaborating on this potentially large wage gap. In equilibrium group R is completely specialized in education A. Therefore the wage and unemployment gap are determined by the trade off that workers from group G are making. Choosing education A gives Green workers a higher wage than education B, but in education B there would be few Green colleagues, and therefore fewer job contacts. Therefore choosing A would result in a lower employment rate for Green workers. What is surprising is that this unemployment gap may be quite small compared to the wage gap that compensates the unemployment gap. In particular, in our simulation, at $\alpha = \hat{\alpha}$, the wage gap of 30 % is compensated by an employment gap of about 3 %. The reason for this tenfold magnification is risk aversion of individuals. Individuals try to avoid the (small) risk of unemployment, in which they have a payoff equal to 0, and they are willing to accept even major losses in income in order to accomplish that.[21]

We would like to know whether an even larger wage gap can be sustained in a partial segregation equilibrium when $\alpha > \hat{\alpha} = .5904$. We therefore plot the equilibrium wages, $w_A(1,\mu^*)$ and $w_B(1,\mu^*)$, and equilibrium employments, $s_A^R(1,\mu^*)$, $s_B^R(1,\mu^*)$, $s_A^G(1,\mu^*)$ and $s_B^G(1,\mu^*)$, as function of α. Remember that the equilibrium μ^* equals zero when $\alpha \leq \hat{\alpha}$, and solves $\Delta\Pi^G(1,\mu^*) = 0$ when $\alpha > \hat{\alpha}$. These plots are shown in Figures 4.2 and 4.3.

The pictures clearly confirm Propositions 4 and 5. Moreover, for the chosen parameters we also observe that the wage gap $G(1,\mu^*)$ is maximized at $\alpha = \hat{\alpha}$. When α becomes

violated. Nonetheless, there is still a unique equilibrium for all values of α.

[21] The risk aversion effect, and thus the wage gap, may be smaller if unemployment is only temporary, and individuals only care about permanent income. On the other hand, from prospect theory it is known that individual agents tend to emphasize small probabilities (Kahneman and Tversky 1979), and thus the small probability of becoming unemployed may get excessive weight in the education decision.

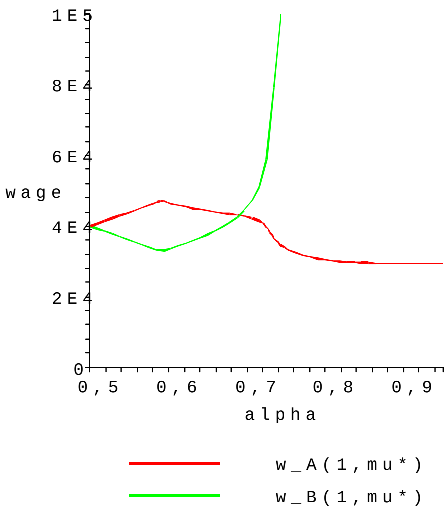

Figure 4.2: Equilibrium wages as function of α.

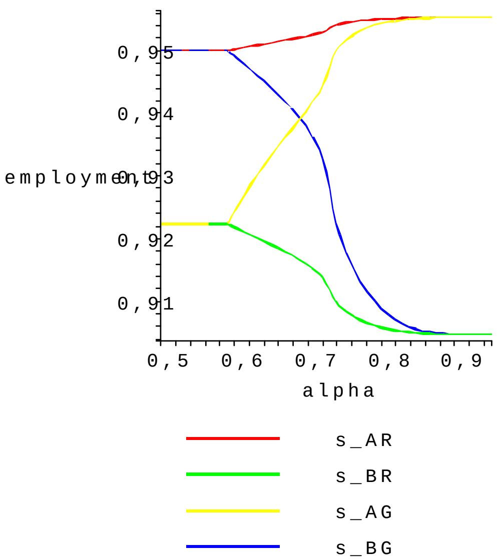

Figure 4.3: Equilibrium employment rates as function of α.

larger than $\hat{\alpha}$, the wage of A declines and the wage of B increases until the wage gap is reversed.

We next look at the sensitivity of $\hat{\alpha}$ with respect to the parameter choices, as we saw that at $\alpha = \hat{\alpha}$ the wage gap is maximized. We do this by computing the elasticities of $\hat{\alpha}$ and of the implied wage gap $G(1, 0)$ at the chosen parameters. That is, we look at the percentage increase of $\hat{\alpha}$ and the maximum wage gap change when a parameter increases by 1% . The elasticities are shown in columns 2 and 3 of Table 4.2. We note that $\hat{\alpha}$ and the implied maximum wage gap are most sensitive to $\rho\theta$, the coefficient of relative risk aversion. A 1% increase in this coefficient leads to a 2% increase in the maximum wage gap. On the other hand, our calibration seems least sensitive to the network parameters $c_1(p + \kappa)$ and $c_1\lambda$. The maximum wage gap seems to be close to linear with respect to $1 - s_0$, the unemployment rate if a worker only consider direct search techniques. That is, if we chose $s_0 = .95$ instead of $s_0 = .90$, it would roughly halve the maximum wage gap.

Implications for the second-best welfare outcome

We now consider the analysis of a second-best optimum. Namely, we suppose that the government (social planner) does not have the institutions to completely control μ_R and μ_G, but that it is able to stabilize a symmetric equilibrium, such that $\mu_R = \mu_G = \mu^S$.[22] Should the government do this? In case the government stabilizes integration, we still impose the equilibrium condition, which is in this case symmetric. Therefore

$$\Pi_A^R(\mu^S, \mu^S) = \Pi_B^R(\mu^S, \mu^S) = \Pi_A^G(\mu^S, \mu^S) = \Pi_B^G(\mu^S, \mu^S).$$

Hence, in the symmetric case there is complete equality. On the other hand, in the case of segregation, we consider the equilibrium allocation $(\mu_R, \mu_G) = (1, \mu^*)$, such that Reds obtain a higher payoff than Greens. Therefore, we might face a tradeoff when assessing an integration policy. It enforces equality, but it might decrease employment.

To this purpose we plot the increase in social welfare from such an integration policy, $I = W(\mu^S, \mu^S)/W(1, \mu^*) - 1$, as function of α. Figure 4.4 shows this plot.

We observe that I is negative for all values of α. So for the chosen parameters the integration policy is never preferred. People are better off segregated.

Our results are very clear; a second best policy involves a "laissez-faire" policy, such that society becomes segregated. The intuition behind this result is twofold. First, in the case of partial segregation the equilibrium is determined by the Green workers. They

[22] In the proof of Lemma 7 we show that there exists a symmetric equilibrium, but that it is unstable; that is, after a small deviation from the equilibrium individual incentives drive education choices to segregation.

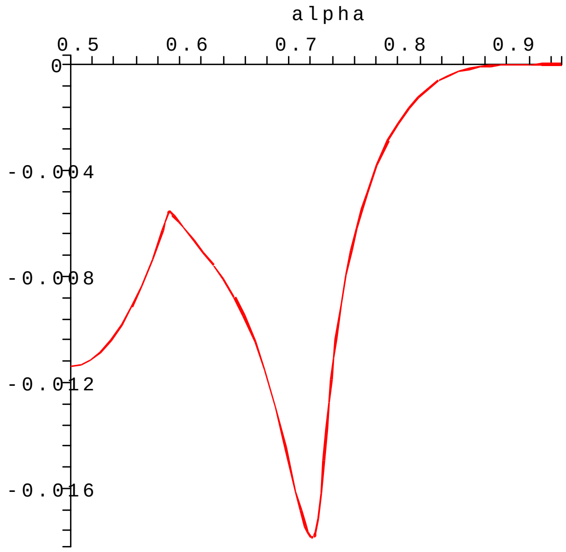

Figure 4.4: The percentage increase in welfare of a policy that enforces perfect integration.

trade off a benefit in wage against a loss in employment. Their individual incentives therefore already put a limit on the amount of wage inequality that can be sustained in equilibrium. Second, an integration policy would lead to lower employment rates. In a society with risk-averse individuals, society puts large emphasis on unemployment, and therefore prefers to allow for some inequality in order to obtain these higher employment rates.

We finally remark that an integration policy is only beneficial when society has *additional* distributional concerns that are not captured by the concavity of the individual utility function. For example, consider the case of a maximin social welfare function: $W_{\min} = \min_i \Pi_i$. In the integrated case, $\mu_R = \mu_G = \mu^S$, everyone obtains the same payoff, whereas in the segregated case workers from group G are worse off. Therefore, $W_{\min}(1, \mu^*) = \Pi_B^G(1, \mu^*)$ and $W_{\min}(\mu^S, \mu^S) = \Pi_B^G(\mu^S, \mu^S)$. We show a comparison of these two payoffs, $\Pi_B^G(\mu^S, \mu^S)/\Pi_B^G(1, \mu^*) - 1$, in Figure 4.5.

We observe that the Green workers would benefit from an integration policy for values of α around $\hat{\alpha}$, where the wage gap is particularly large. In such a case, strong distributional concerns would justify integration.

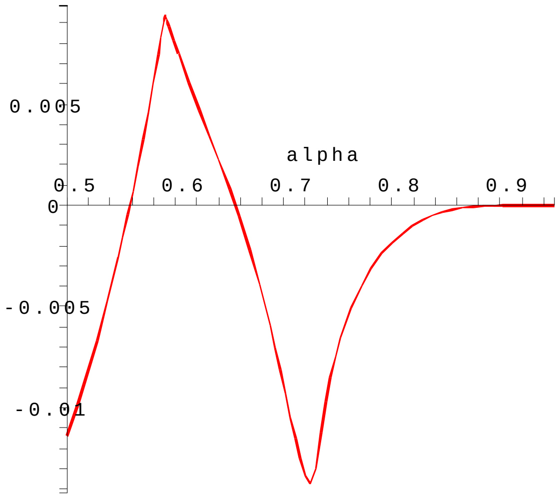

Figure 4.5: The percentage increase in payoffs for Green workers of a policy that enforces perfect integration.

4.6 Summary and conclusions

We have investigated a simple social network framework where jobs are obtained through a network of contacts formed stochastically, after career decisions had been made. We have established that even with a very small amount of homophily within each social group, stable occupational segregation equilibria will arise. If the wage differential across the occupations is not too large, complete segregation will always be sustainable. If the wage differential is large, complete segregation cannot be sustained, but a partial segregation equilibrium in which one of the group fully specializes in one education while the other group mixes over the career tracks, is sustainable. Furthermore, our model is able to explain sustained unemployment and wage differences between the social groups.

We also analyze the implications of our model from a social planner's point of view. In the first best social welfare optimum, we find that segregation is the socially preferred outcome. Subject to proper calibration of our model parameters, a second best social welfare analysis supports a laissez-faire policy, where society also becomes segregated, shaped by individual incentives. Both these conclusions are valid in light of 'reasonable' concavity features of the individual utility function. Our social welfare conclusions cast some doubts on an "always integration" policy choice; if job referrals through contact networks are relevant in matching workers to vacancies, and if the mechanisms of our model are the correct ones, an integration approach would only be justified under strong additional distributional concerns, not reflected in the individual utility functions.

We assumed that individuals first choose an education, and then form a network of job contacts. As a consequence, individuals have to make *expectations* about the network they could form, and base their education decisions on these expectations. This is in contrast to earlier work on the role of networks in the labor market. In former research, the network was supposed to be already in place, or the network was formed in the first stage (Montgomery 1991, Calvó-Armengol 2004, Calvó-Armengol and Jackson 2004).

Our departure from the earlier frameworks raises questions about the assumed timing of the education choice. Are crucial career decisions made before or after job contacts are formed? One might be tempted to answer: both. Of course everyone is born with family ties, and in early school and in the neighborhood children form more ties. It is also known that peer-group pressure among children has a strong effect on decisions to, for instance, smoke or engage in criminal activities and, no doubt, family and early friends do form a non-negligible source of influence when making crucial career decisions. However, we argue that most *job-relevant contacts* (the so called 'instrumental ties') are made later, for instance at the university, or early at the workplace, hence after a specialized career track had been chosen. In spite of the fact that those ties are typically not as strong as family ties, they are more likely to provide relevant information on vacancies to job seekers; Granovetter (1973, 1995) provides convincing evidence that job seekers more often receive crucial job information from acquaintances ("weak ties"), rather than from family or very close friends ("strong ties"). If the vast majority of such instrumental ties are formed after the individual embarked on a (irreversible) career, then it is justified to consider a model in which the job contact network is formed after making a career choice.

While our social interaction model can describe empirical patterns of occupational segregation and wage inequality between gender, racial or ethnical groups, other factors are also documented to play a significant role in this context. This model should thus be seen as complement to alternatives, such as taste discrimination or rational bias by employers, which are still present in the market despite their (predicted) erosion over time, due to both competitive pressure and institutional instruments. It is therefore pertinent to directly investigate in future research how relevant are the mechanisms described in this paper and to assess their relative strength in explaining observed occupational segregation, vis-à-vis other proposed theories.

Our model easily allows for interesting extensions. One avenue for future research is to extend our framework to issues such as the position of minority versus majority groups, by modeling the interaction between social groups of unequal sizes. Another avenue is to consider heterogeneity in productivity. This would allow us to analyze the mismatch of workers to firms due to network effects. We intend to pursue these lines of research in the future.

Appendix Chapter 4: Proofs

4.A Proofs for all propositions

The proof of Proposition 3 uses the following lemma:

Lemma 7 *Suppose Assumptions 1 and 2 hold. A weakly stable equilibrium* (μ_R^*, μ_G^*)*, in which* $0 < \mu_R^* < 1$ *and* $0 < \mu_G^* < 1$*, does not exist.*

Proof. Suppose (μ_R^*, μ_G^*) is a stable equilibrium, and $\mu_R^* \in (0,1)$ and $\mu_G^* \in (0,1)$. By condition (4.2)

$$\Pi_A^R(\mu_R^*, \mu_G^*) = \Pi_B^R(\mu_R^*, \mu_G^*) \text{ and } \Pi_A^G(\mu_R^*, \mu_G^*) = \Pi_B^G(\mu_R^*, \mu_G^*) \tag{4.18}$$

Substituting (4.8)-(4.9) into (4.18) and rewriting, these equations become

$$\frac{U(w_A(\mu_R^*, \mu_G^*))}{U(w_B((\mu_R^*, \mu_G^*))} = \frac{s_B^R(\mu_R^*, \mu_G^*)}{s_A^R(\mu_R^*, \mu_G^*)} = \frac{s_B^G(\mu_R^*, \mu_G^*)}{s_A^G(\mu_R^*, \mu_G^*)}. \tag{4.19}$$

Since $\lambda > 0$, $\mu_R^* > \mu_G^*$ implies $s_A^R > s_A^G$ and $s_B^R < s_B^G$. But this means that if $\mu_R^* > \mu_G^*$, then

$$\frac{s_B^R(\mu_R^*, \mu_G^*)}{s_A^R(\mu_R^*, \mu_G^*)} < \frac{s_B^G(\mu_R^*, \mu_G^*)}{s_A^G(\mu_R^*, \mu_G^*)}.$$

which contradicts (4.19). The same reasoning holds for $\mu_R^* < \mu_G^*$. Hence, it must be that $\mu_R^* = \mu_G^*$.

However (μ_R^*, μ_G^*) with $\mu_R^* = \mu_G^*$ cannot be a stable equilibrium. To see this, suppose that (μ^*, μ^*) with $\mu^* \in (0,1)$ is a stable equilibrium. Hence $\Pi_A^X(\mu^*, \mu^*) = \Pi_B^X(\mu^*, \mu^*)$ for $X \in \{R, G\}$ and $\frac{\partial \Delta \Pi^X}{\partial \mu_X} < 0$ at $\mu_R = \mu_G = \mu^*$, and $\det(G(\mu^*, \mu^*) > 0$, where $G(\mu) = D\Delta\Pi(\mu)$ is the Jacobian of $\Delta\Pi \equiv (\Delta\Pi^R, \Delta\Pi^G)$ with respect to $\mu \equiv (\mu_R, \mu_G)$.

Since $\lambda > 0$, it must be that

$$\frac{\partial s_A^X}{\partial \mu_X} > \frac{\partial s_A^X}{\partial \mu_Y} > 0 \tag{4.20}$$

and

$$\frac{\partial s_B^X}{\partial \mu_X} < \frac{\partial s_B^X}{\partial \mu_Y} < 0 \tag{4.21}$$

for $X, Y \in \{R, G\}$ and $Y \neq X$. Furthermore, if $\mu_R = \mu_G = \mu^*$, then $s_A^X = s_A^Y$, $\frac{\partial L_A}{\partial \mu_X} = \frac{\partial L_A}{\partial \mu_Y}$, $\frac{\partial L_B}{\partial \mu_X} = \frac{\partial L_B}{\partial \mu_Y}$, and therefore,

$$\frac{\partial w_A}{\partial \mu_X} = \frac{\partial w_A}{\partial \mu_Y} \tag{4.22}$$

and

$$\frac{\partial w_B}{\partial \mu_X} = \frac{\partial w_B}{\partial \mu_Y}. \tag{4.23}$$

From (4.20)-(4.23) and Assumption 2, it follows that, at $\mu_R = \mu_G = \mu^*$,

$$\frac{\partial \Delta \Pi^X}{\partial \mu_Y} < \frac{\partial \Delta \Pi^X}{\partial \mu_X} < 0.$$

for $X, Y \in \{R, G\}$, $X \neq Y$. But then it is straightforward to see that $\det(G(\mu^*, \mu^*)) < 0$. This contradicts stability. ■

Proof of Proposition 3. (i) If (4.10) holds, then

$$\Pi_A^R(1,0) > \Pi_B^R(1,0) \text{ and } \Pi_A^G(1,0) < \Pi_B^G(1,0).$$

Hence, $(\mu_R, \mu_G) = (1, 0)$ is clearly a stable equilibrium. The same is true for $(\mu_R, \mu_G) = (0, 1)$. Lemma 7 and Assumption 2 ensure that these are the only two equilibria.

(ii) If (4.11) is true, then

$$\Pi_A^G(1,0) > \Pi_B^G(1,0). \tag{4.24}$$

Furthermore, from Assumption 1 we know that $\frac{\partial \Delta \Pi^G(1,\mu_G)}{\partial \mu_G} < 0$ for all $\mu_G \in [0, 1]$. It follows from Assumption 1, equation (4.24) and continuity of F, U and s, that there must be a unique μ^*, such that

$$\Pi_A^G(1,\mu^*) = \Pi_B^G(1,\mu^*).$$

Moreover, $s_A^R(1,\mu^*) > s_A^G(1,\mu^*)$ and $s_B^G(1,\mu^*) > s_B^R(1,\mu^*)$, so we have at $(\mu_R, \mu_G) = (1, \mu^*)$

$$\Pi_A^X > \Pi_B^Y = \Pi_A^Y > \Pi_B^X. \tag{4.25}$$

It is therefore clear that $(\mu_R, \mu_G) = (1, \mu^*)$ is a stable equilibrium. The same is true for $(\mu_R, \mu_G) = (\mu^*, 1)$.

To show that there is no other equilibrium, note that by (4.11) $\Pi_A^R(1,0) > \Pi_B^R(1,0)$. Assumption 2 then implies that $\Pi_A^R(\mu,0) > \Pi_B^R(\mu,0)$ for all $\mu \in [0,1]$. Hence, $(\mu, 0)$ and, similarly, $(0, \mu)$ cannot be an equilibrium. By Lemma 7 we also know that there is no mixed equilibrium. ■

Proof of Proposition 4. The equations follow almost directly. We have

$$s_A^R(1,0) = s_B^G(1,0) = s_H > s_L = s_B^R(1,0) = s_A^G(1,0).$$

Further, by assumption $w_A \geq w_B$ at $(\mu_R, \mu_G) = (1, 0)$. Finally, at $(\mu_R, \mu_G) = (1, 0)$

$$U(w_A)s_A^R \geq U(w_B)s_B^G \geq U(w_A)s_A^G \geq U(w_B)s_B^R,$$

and this is equivalent to (4.12). ■

Proof of Proposition 5. Consider the stable equilibrium at $(1, \mu^*)$. Since it is an equilibrium we know that

$$\Pi_A^G(1, \mu^*) = \Pi_B^G(1, \mu^*).$$

In the proof of Proposition 3, equation (4.25), we already demonstrated the inequality (4.14) Further, by Assumption 2 we know that $\Delta\Pi^G(1, \mu_G)$ is strictly monotonically decreasing in μ_G.

(i) If $\hat{\mu} < \frac{\lambda}{2(p+\kappa+\lambda)}$, then $s_A^G(1, \hat{\mu}) < s_B^G(1, \hat{\mu})$. As $w_A(1, \hat{\mu}) = w_B(1, \hat{\mu})$ it must be that

$$\Pi_A^G(1, \hat{\mu}) < \Pi_B^G(1, \hat{\mu}).$$

But then it also must be that $\mu^* < \hat{\mu}$. As we consider a partial equilibrium, we know that $\mu^* > 0$. Hence, $0 < \mu^* < \hat{\mu}$ and $w_A(1, \hat{\mu^*}) > w_B(1, \hat{\mu^*})$, as $w_A(\mu_R, \mu_G)$ is a decreasing function, whereas $w_B(\mu_R, \mu_G)$ is increasing.

(ii) If $\hat{\mu} > \frac{\lambda}{2(p+\kappa+\lambda)}$, then $s_A^G(1, \hat{\mu}) > s_B^G(1, \hat{\mu})$ and $\Pi_A^G(1, \hat{\mu}) < \Pi_B^G(1, \hat{\mu})$. But then $\mu^* > \hat{\mu}$. By Assumption 1 we know that $\mu^* < 1$. Hence, $\hat{\mu} < \mu^* < 1$, and therefore $w_A(1, \hat{\mu^*}) < w_B(1, \hat{\mu^*})$ ■

We next continue with the proof of Proposition 6. This proof uses the following lemma:

Lemma 8 *Suppose that for all $x \in [0, (p + \kappa + \lambda)/2]$*

$$s''(x) > -\frac{4}{\lambda}\, s'(x). \tag{4.26}$$

(i) If $\mu_X > \mu_Y$ for $X, Y \in \{R, G\}$, then

$$\frac{\partial L_A}{\partial \mu_X}(\mu_R, \mu_G) > \frac{\partial L_A}{\partial \mu_Y}(\mu_R, \mu_G) > 0, \tag{4.27}$$

and

$$\frac{\partial L_B}{\partial \mu_Y}(\mu_R, \mu_G) < \frac{\partial L_B}{\partial \mu_X}(\mu_R, \mu_G) < 0. \tag{4.28}$$

(ii) If $\mu_R = \mu_G = \mu$, then

$$\frac{\partial^2 L_A}{(\partial \mu_X)^2}(\mu, \mu) > \frac{\partial^2 L_A}{\partial \mu_X \partial \mu_Y}(\mu, \mu), \tag{4.29}$$

and

$$\frac{\partial^2 L_B}{(\partial \mu_X)^2}(\mu,\mu) > \frac{\partial^2 L_B}{\partial \mu_X \partial \mu_Y}(\mu,\mu). \tag{4.30}$$

Proof. (i) It is easy to derive that for $X \in \{R, G\}$:

$$\frac{\partial L_A}{\partial \mu_X} = \frac{1}{2}\left(s_A^X + \mu_R \frac{\partial s_A^R}{\partial \mu_X} + \mu_G \frac{\partial s_A^G}{\partial \mu_X}\right) > 0 \tag{4.31}$$

$$\frac{\partial L_B}{\partial \mu_X} = \frac{1}{2}\left(-s_B^X + (1-\mu_R)\frac{\partial s_B^R}{\partial \mu_X} + (1-\mu_G)\frac{\partial s_B^G}{\partial \mu_X}\right) < 0 \tag{4.32}$$

at (μ_R, μ_G). From (4.31) and (4.32), we find that for all $X, Y \in \{R, G\} : \partial L_A/\partial \mu_X > \partial L_A/\partial \mu_Y$ is equivalent to

$$s_A^X + \mu_X\left(\frac{\partial s_A^X}{\partial \mu_X} - \frac{\partial s_A^X}{\partial \mu_Y}\right) > s_A^Y + \mu_Y\left(\frac{\partial s_A^Y}{\partial \mu_Y} - \frac{\partial s_A^Y}{\partial \mu_X}\right). \tag{4.33}$$

With the definition of s_A^X in (4.4) we can write out

$$s_A^X + \mu_X\left(\frac{\partial s_A^X}{\partial \mu_X} - \frac{\partial s_A^X}{\partial \mu_Y}\right) = s\left((p+\kappa)\bar{\mu} + \lambda\mu_X/2\right) + \frac{\mu_X \lambda}{2} s'\left((p+\kappa)\bar{\mu} + \lambda\mu_X/2\right) \tag{4.34}$$

when $X \neq Y$. Therefore $\mu_X > \mu_Y$ is equivalent to (4.33), whenever (4.34) is strictly monotone increasing with μ_X, where we can treat $\bar{\mu} = (\mu_X + \mu_Y)/2$ as a constant. It is easy to check that this is indeed the case under condition (4.26). We conclude that hypothesis (4.27) holds whenever $\mu_X > \mu_Y$. With a similar derivation one can show that condition (4.26) implies (4.28) as well.

(ii) The second derivatives of L_A and L_B with respect to μ_X and μ_Y are

$$\frac{\partial^2 L_A}{\partial \mu_X \partial \mu_Y} = \frac{1}{2}\left(\frac{\partial s_A^X}{\partial \mu_Y} + \frac{\partial s_A^Y}{\partial \mu_X} + \mu_R \frac{\partial^2 s_A^R}{\partial \mu_X \partial \mu_Y} + \mu_G \frac{\partial^2 s_A^G}{\partial \mu_X \partial \mu_Y}\right) \tag{4.35}$$

$$\frac{\partial^2 L_B}{\partial \mu_X \partial \mu_Y} = \frac{1}{2}\left(-\frac{\partial s_B^X}{\partial \mu_Y} - \frac{\partial s_B^Y}{\partial \mu_X} + (1-\mu_R)\frac{\partial^2 s_B^R}{\partial \mu_X \partial \mu_Y} + (1-\mu_G)\frac{\partial^2 s_B^G}{\partial \mu_X \partial \mu_Y}\right). \tag{4.36}$$

Taking the second derivatives of s_A^X, evaluating at $\mu_R = \mu_G = \mu$ and reordering, we get that (4.29) is equivalent to

$$s''((p+\kappa+\lambda)\mu/2) < -\frac{4}{\lambda\mu} s'((p+\kappa+\lambda)\mu/2). \tag{4.37}$$

Inequality (4.37) clearly holds if condition (4.26) holds, which proves (4.29). In a similar fashion, (4.26) implies (4.30) ■

Proof of Proposition 6. Suppose that $W(\mu_R, \mu_G)$ is maximized at $(\mu_R, \mu_G) = (\tilde{\mu}_R, \tilde{\mu}_G)$, where $\tilde{\mu}_R \in (0,1)$ and $\tilde{\mu}_G \in (0,1)$. Define $c \equiv L_A(\tilde{\mu}_R, \tilde{\mu}_G)/L_B(\tilde{\mu}_R, \tilde{\mu}_G)$, and consider the constrained maximization problem:

$$\max_{\mu_R \in [0,1], \mu_G \in [0,1]} W(\mu_R, \mu_G) \text{ s.t. } L_A(\mu_R, \mu_G) = cL_B(\mu_R, \mu_G). \tag{4.38}$$

Because by definition of c, the solution $(\tilde{\mu}_R, \tilde{\mu}_G)$ satisfies the restriction

$$g(\mu_R, \mu_G) = cL_B(\mu_R, \mu_G) - L_A(\mu_R, \mu_G) = 0, \tag{4.39}$$

it actually solves the maximization problem (4.38).

Define the feasible set $C = \{\mu_R \in [0,1], \mu_G \in [0,1] | g(\mu_R, \mu_G) = 0\}$. By the assumption of constant returns to scale, we have that for all $(\mu_R, \mu_G) \in C$: $w_A(\mu_R, \mu_G)$ and $w_B(\mu_R, \mu_G)$ are constant, and therefore, at all $(\mu_R, \mu_G) \in C$, the welfare function (4.16) can be written as

$$W(\mu_R, \mu_G) = L_A(\mu_R, \mu_G)(U(\overline{w_A}) + U(\overline{w_B})/c),$$

which is monotone increasing with $L_A(\mu_R, \mu_G)$. Therefore, the solution $(\tilde{\mu}_R, \tilde{\mu}_G)$ also solves the following maximization problem:

$$\max_{\mu_R \in [0,1], \mu_G \in [0,1]} L_A(\mu_R, \mu_G) \text{ s.t. } L_A(\mu_R, \mu_G) = cL_B(\mu_R, \mu_G). \tag{4.40}$$

We verify that $(\tilde{\mu}_R, \tilde{\mu}_G)$ indeed satisfy the first- and second-order conditions of problem (4.40). The Lagrangian is given by

$$\mathcal{L}(\mu_R, \mu_G, \psi) = (1 - \psi)L_A(\mu_R, \mu_G) + \psi c L_B(\mu_R, \mu_G).$$

Since $(\tilde{\mu}_R, \tilde{\mu}_G)$ is supposed to be interior, the following first order constraints should hold:

$$\frac{\partial \mathcal{L}}{\partial \mu_R}(\tilde{\mu}_R, \tilde{\mu}_G, \psi) = (1 - \psi)\frac{\partial L_A}{\partial \mu_R}(\tilde{\mu}_R, \tilde{\mu}_G) + \psi \frac{\partial L_B}{\partial \mu_R}(\tilde{\mu}_R, \tilde{\mu}_G) = 0 \tag{4.41}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_G}(\tilde{\mu}_R, \tilde{\mu}_G, \psi) = (1 - \psi)\frac{\partial L_A}{\partial \mu_G}(\tilde{\mu}_R, \tilde{\mu}_G) + \psi \frac{\partial L_B}{\partial \mu_G}(\tilde{\mu}_R, \tilde{\mu}_G) = 0. \tag{4.42}$$

The first part of Lemma 8 implies that $\psi \in (0,1)$ and that under condition (4.26): $\mu_R > \mu_G$ if and only if $\partial \mathcal{L}/\partial \mu_R > \partial \mathcal{L}/\partial \mu_G$. Therefore, condition (4.26) and the first-order conditions imply that $\tilde{\mu}_R = \tilde{\mu}_G \equiv \tilde{\mu}$.

Since $\tilde{\mu}_R = \tilde{\mu}_G$ defines a unique point in C, the second-order condition should hold at $\tilde{\mu}_R = \tilde{\mu}_G$, which says that the Hessian of the Lagrangian with respect to (μ_R, μ_G)

evaluated at the social optimum, $D^2_{\mu_R,\mu_G}\mathcal{L}(\tilde{\mu},\tilde{\mu},\psi)$, is negative definite on the subspace $\{z_R, z_G | z_R(\partial g/\partial \mu_R) + z_G(\partial g/\partial \mu_G) = 0\}$. The second order condition is thus that at $(\mu_R,\mu_G) = (\tilde{\mu},\tilde{\mu})$:

$$2\frac{\partial g}{\partial \mu_R}\frac{\partial g}{\partial \mu_G}\frac{\partial^2 \mathcal{L}}{\partial \mu_R \partial \mu_G} - \left(\frac{\partial g}{\partial \mu_R}\right)^2 \frac{\partial^2 \mathcal{L}}{(\partial \mu_G)^2} - \left(\frac{\partial g}{\partial \mu_G}\right)^2 \frac{\partial^2 \mathcal{L}}{(\partial \mu_R)^2} > 0. \tag{4.43}$$

Because $\frac{\partial g}{\partial \mu_R}(\tilde{\mu},\tilde{\mu}) = \frac{\partial g}{\partial \mu_G}(\tilde{\mu},\tilde{\mu})$, and $\frac{\partial^2 \mathcal{L}}{(\partial \mu_G)^2}(\tilde{\mu},\tilde{\mu}) = \frac{\partial^2 \mathcal{L}}{(\partial \mu_R)^2}(\tilde{\mu},\tilde{\mu})$, the second order condition (4.43) simplifies to $\frac{\partial^2 \mathcal{L}}{\partial \mu_R \partial \mu_G}(\tilde{\mu},\tilde{\mu}) > \frac{\partial^2 \mathcal{L}}{(\partial \mu_R)^2}(\tilde{\mu},\tilde{\mu})$, or equivalently

$$(1-\psi)\frac{\partial^2 L_A}{\partial \mu_R \partial \mu_G}(\tilde{\mu},\tilde{\mu}) + \psi\frac{\partial^2 L_B}{\partial \mu_R \partial \mu_G}(\tilde{\mu},\tilde{\mu}) > (1-\psi)\frac{\partial^2 L_A}{(\partial \mu_R)^2}(\tilde{\mu},\tilde{\mu}) + \psi\frac{\partial^2 L_B}{(\partial \mu_R)^2}(\tilde{\mu},\tilde{\mu}). \tag{4.44}$$

By the second part of Lemma 8, inequality (4.44) cannot hold under condition (4.26). Therefore we have a contradiction and the non-segregation allocation $(\tilde{\mu}_R, \tilde{\mu}_G)$ cannot be a social optimum. Since a social optimum exists by continuity of W and compactness of $[0,1]^2$, the social optimum necessarily has to involve complete or partial segregation.

Chapter 5

The impact of workplace conditions on firm performance

5.1 Introduction

In this paper we investigate which firm characteristics associate with good work environment practice and the impact of workplace conditions on firm performance. Despite a sizable economic literature that has paid attention to determinants of capital investments, technological innovations or work reorganization in general, and to their respective effect on firm or establishment financial performance, there has been virtually no study on the impact of detailed, physical as well as psycho-social, work environment health and safety conditions, on firm performance indicators. Ours is the first study to focus on the effects of specific health and safety workplace indicators on firm productivity and mean wage. We are able to link detailed work conditions data from a representative Danish cross-sectional survey of establishments to the longitudinal register matched employer-employee data, merged with information on the firms' business accounts. This allows us to use empirical specifications where we can address to a considerable extent econometric problems typical in such contexts, such as omitted variables or endogeneity.

Work environment related issues have been prioritized in labour policy debates all

[0]This chapter is based on Buhai, Cottini and Westergård-Nielsen (2008). We are grateful for comments and suggestions to Paul Bingley, Nick Bloom, Harald Dale-Olsen, Tor Eriksson, Nathalie Greenan, Lisa Lynch, John van Reenen, Michael Waldman, Thomas Zwick, and to participants in seminars and conferences at CAED in Budapest, SOLE in New York, Industrial Relations Center of the Carlson School of Management-University of Minnesota, SMYE in Lille, Center for Corporate Performance Workshop in Ebeltoft, and National Institute for Occupational Health (AMI) in Copenhagen. AMI is the provider of the work environment data used in this project. We also thank Kenneth Sørensen and Philip Røpcke for help with handling the data. The usual disclaimers apply.

throughout the industrialized nations. Improving the general work environment has been for instance a declared target of the European Union, as stated in the consolidated version of the Treaty establishing the European Community.[1] More recently, the 2001 report on employment of the European Commission includes specific work conditions in its "social policy agenda".[2] The same EC report concludes by stating that– although "job quality" is acknowledged to have generally improved within the EU– "working conditions" are still an exception; for instance, the total costs of occupation-related health risks and accidents are estimated to be enormous, with values in the range of 2.5%- 4% of the EU member states' GNPs[3]. The estimated costs of job-related illnesses for the USA are equally large, cca. 3% GNP, see e.g. Leigh et al (1996). See also Figure 2 in the next section for a per-country histogram of estimated aggregate costs of job-related risks and illnesses.

Despite the hot policy context, intuitive implications of the macro-level discussion mentioned above have been so far neither backed up, nor falsified by thorough empirical research using microdata. We do not know for instance whether in practice a "better workplace environment" actually pays off in terms of higher worker productivity or, for that matter, to what extent "bad" workplace health and safety conditions are compensated for by wage premia. Our paper aims to help in filling this knowledge gap and contribute to the research based evidence in the microeconomics of the firm's work environment and production organization. Thus, we believe it is important to know both i) which firm and aggregate employee characteristics are statistically associated with better workplace conditions and, crucially, ii) the impact of enacting/improving specific work environment conditions on the performance of firms. To give a concrete example for i), do written work environment rules or work environment training courses for all employees, but also, e.g., higher aggregate human capital level, proportion of managers, female employees in the firm, age of the firm etc., associate with better workplace environment quality? At the same time, expenditures by firms to improve workplace conditions should be seen as investments in the economic sense, i.e. costs borne today in order to reap benefits in terms of higher profits tomorrow. Such investment decisions from the part of the employer

[1] In the protocol of the Treaty of Maastricht (1992), the social compentencies of the European Community were expanded to include "working conditions". A "European Foundation of The Improvement of Living and Working Conditions" had been in place already since 1975.

[2] Explicit reference is made to: intrinsic job quality; skills, life-long learning, and career development; gender equality; health and safety at work; flexibility and security; inclusion and access to the labour market; work organisation and work-life balance; social dialogue and worker involvement; diversity and non-discrimination; overall work performance.

[3] Citing directly from the text of the report, "The evolution of job quality in the EU in recent years was generally positive, with the exception of working conditions which do not seem to have improved. Accidents at the workplace and occupational diseases remain a challenge to the EU economies, with direct and indirect costs due to work-related health risks and accidents at work estimated to amount to between 2.6% and 3.8% of GNP in the EU".

need therefore to be strategic; it is not ex ante obvious which of the specific dimensions of the workplace environment should be targeted, and in which way an improvement in them would impact firm productivity or employee welfare. Hence, to consider an example for ii), should one pay equal attention to perceived physical workplace problems such as noise or heavy lifting burden or internal climate conditions, and to perceived problems in the psycho-social realm (decision lattitude of the employees, stress, working with colleagues etc.)? Are these workplace environment dimensions equally relevant in enhancing firm productivity and/or should they be equally compensated for by higher wages when unsolved? The empirical literature so far has indeed been silent[4] on whether better workplace environment –and if so, precisely which dimensions of the "workplace environment"– leads to a better firm productivity, and whether workplaces where work environment is perceived more hazardous than in others are more likely to pay employees a job hazard premium. A priori, one can for instance envisage at least two channels through which good health and safety conditions at the workplace could be improving firm performance: on the one hand, the employee pool would likely be more satisfied/enthusiastic and hence directly more productive at the job and/or the firm would be more able to retain the most productive employees[5], while on the other hand, there will be less problems related to absenteeism due to job-related illnesses and diseases, which again might indirectly translate in better firm performance. As stated earlier however, it is ultimately an (so far, unanswered) empirical question whether in practice the reasoning above is confirmed and if so, to what extent; i.e. whether improvement in all, or perhaps only in some of the specific workplace conditions implies higher marginal firm productivity.

To the best of our knowledge, there have not been so far any studies explicitly analysing determinants of workplace health and safety conditions or the impact of such workplace practice on firm productivity and/or wages, in country-wide representative datasets. The few studies that come somewhat close to ours in terms of focus, though only indirectly address our concerns, are case studies such as Katz et al (1983), who analyse the relationship among plant–level measures of industrial relations performance, economic performance and quality of working life programs, among plants within a division of General Motors, or Gittel et al (2004), who investigate the link between quality of labor relations

[4] A legitimate concern would also be the precise theoretical connection between workplace environment conditions and firm performance. While this has not been modeled explicitly in the existent literature, what we have in mind here is a similar mechanism as that between various (general) organizational change proxies and firm performance, obviously inheriting all analogous problems related to endogeneity and reverse causality.

[5] Ample evidence showing that employee attitudes influenced by workplace organization can have significant efects on economic outcomes appear in several papers. One such recent study is for instance Bartel et al (2003).

(understood as union representation), shared governance, wages and firm performance, in the airline industry. More generally, there is also a large, ongoing, literature focusing on the impact of firms' industrial resource management system and general reorganization therein, on firm financial performance; e.g., a number of recent studies conclude by promoting the advantages of using high involvement or high commitment human resources practices (e.g. Osterman, 1994; Gittleman et al. 1998 and Batt, 2002). A few other studies have found empirical links between the use of such practices and overall firm-level performance (e.g. Huselid, 1995; Osterman, 2000; Cappelli and Newmark, 2001; Caroli and Van Reenen, 2001; Guthrie 2001), while yet others have gone in more detail, but narrowed the scope of their analysis to particular industries (Batt, 1999; Ichiniowski et al., 1997; Ichiniowski and Shaw, 1998). Finally, a number of recent papers have used individual worker data to study the relationship between new workplace practices and workplace safety and health (Askenazy, 2001; Brenner et al., 2004; Askenazy and Caroli, 2006).

In terms of research methodology, Black and Lynch (2001) is the most related study to our paper; they estimate an augmented production function that incorporates variables reflecting work reorganization and firm specific aggregate employee characteristics, next to classical production inputs. While Black and Lynch apply their methodology to investigate workplace reorganization affecting firm productivity, we adapt it for specific improvement in workplace environment health and safety indicators, looking at effects both on firm productivity and on the firms' mean wages[6]. As in Black and Lynch (2001), we have survey data for the workplace environment explanatory variables and independently measured, objective, further firm-specific explanatory and explained variables.

The first part of the empirical analysis consists in estimating binary outcome (logit) models of general and specific work environment quality indicators on several aggregate employee characteristics, as well as on proxies of good practice in terms of work environment, such as e.g. having written work environment rules or offering work environment training courses for all employees. This gives an idea of which such variables are mostly associated with good work environment outcomes, e.g. in the spirit of Osterman (1994), who looked at the association between firm characteristics and human resource reorganization. The second, and main, part of our analysis consists in estimating standard Cobb-Douglas production functions, augmented with employees' aggregated characteristics such as e.g. proportion of females, proportion of unskilled workers, average human capital in the firm, and the specific work environment indicators. The longitudinal di-

[6]Another recent study that succesfully applies the methodology in Black and Lynch, to study the productivity impact of shop-floor employee involvement, is Zwick (2004).

mension of the register firm data enables us to estimate these augmented production functions in two simple steps, using either fixed firm effects (FE) or system-generalized method of moments (GMM) estimations in the first stage, where we only work with the production inputs and aggregate employees' characteristics, and ordinary least squares (OLS) of the mean residuals resulting from the first stage on the cross-sectional work environment indicators, in a second stage. This closely follows the strategy set out in Black and Lynch (2001), allowing us to address eventual endogeneity biases due to unobserved time-invariant firm heterogeneity and simultaneity of classical inputs and output in the production function. Analogous to the estimation of the production functions, we also investigate the explanatory power of work environment conditions and other employee aggregate characteristics in accounting for between-firm mean wage differentials, using firm fixed effects estimation in a first stage, and a second stage that uses the average residual from the first stage regressed on the workplace condition indicators. A major improvement relative to Black and Lynch (2001) is that in our dataset we observe all firm and employee characteristics over time, and not only the evolution of the firm production inputs, and that we can also proxy for likely time-variant unobservables such as managerial ability, which might otherwise remain correlated with the work condition indicators in the second stage OLS estimation, by instrumenting for changes and lagged levels of the proportion of managerial positions over time.

The main findings of our study can be summarized as follows. In terms of firm characteristics associated with good work environment outcomes, the following factors are found to have explanatory power in accounting for the variation in the workplace conditions among firms: the proportion of managerial positions, all-employee work environment courses offered in the firm and, to less extent, the proportion of female employees in the firm's workforce and prioritizing work environment practice at the firm. These variables are statistically significant and of expected signs for several of the specific workplace environment indicators. More important, in terms of effects of work environment indicators on firm performance, our results suggest that only improvement in some of the physical dimensions of workplace environment, specifically "internal climate" and respectively, "repetitive and strenuous work activity" (positively) impacts the firm aggregate productivity. At the same time, the only workplace health and safety condition with explanatory power in the between-firm mean wage differential is the "internal climate", suggesting a compensating wage differential story.

The remainder of the paper is organised as follows. The data and Danish institutional context are overviewed in the following section. In Section 3. we put forward the empirical specification and estimation results for determinants of good workplace conditions.

Section 4 contains the main analysis, the impact of the workplace environment on firm performance, both in terms of firm productivity and firm aggregate wage. Section 5 briefly summarizes and presents some concluding remarks.

5.2 Data description and the Danish context

5.2.1 Denmark and workplace conditions

Studying Denmark in the workplace environment context turns out a very sensible thing to do. First, Denmark tops the OECD charts on job satisfaction of employees with their work conditions, as shown in Figure 1, reproduced from the online statistics source on job quality of the "Canadian Policy Research Networks"[7]. At the same time, Denmark is a country with a very generous social safety net (and publicly funded universal health care system) and might thus be argued to be very vulnerable to externalization of the costs of occupational-related risks/injuries from the employer to the society[8]. Dorman (2000) states for instance that "[i]ronically countries with highly developed public welfare programs are more vulnerable to cost externalization, since these programs either pool risks (dissipating the risk to the individual enterprise) or transfer a portion of the burden to taxpayers. An example would be publicly funded health care systems, which absorb much of the cost of occupational accidents and diseases". However, in terms of estimated total costs ("aggregate economic costs") of occupational-related injury and disease, although these are very high in absolute terms, Denmark does not fare too badly in comparison to other OECD countries– and in particular relative to its Scandinavian neighbours– as seen from Figure 5.2 below, reproduced from Beatson and Coleman (1997), with the US estimate from Leigh et al (1996). Finally, a huge deal of attention has been given and continues to be given to enhancing workplace conditions in Denmark, on the policy stage. For instance, explicit targeting of improvement in both psychosocial and physical workplace conditions has been recently topping the agenda of both the Danish Ministry of Labour and the Danish Working Environment Authority[9], see also Hasle and Moller

[7] The exact web address is http://www.jobquality.ca/indicators/international/satisfaction_main.shtml It is noted that the data for Europe comes from the Third European Study on Working Conditions (2000), while the data for US and Canada comes from the Ekos Rethinking North American Integration Survey (2002).

[8] We are not aware of attempts to decompose the burden of the job-related injury and disease costs on shares of various societal agents for other countries than the US, where Leigh et al (1996) estimate that, out of the approx. 3% of the GDP that is translated in such costs, 11% falls on the employer, 9% on the consumer and 80% on the worker.

[9] The Working Environment Act (1999) introduces for instance several concrete measures aimed at improving the workplace environment, e.g. unannounced screening of all Danish enterprises within a

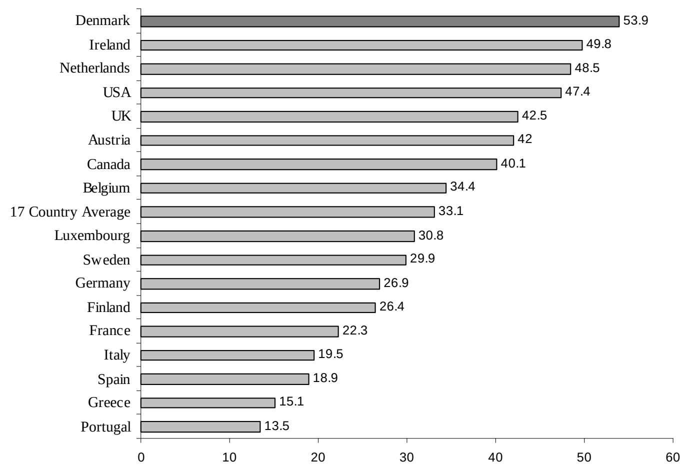

Figure 5.1: Percentage of workers that report being "very satisfied" with working conditions in their main paid job, by country

(2001).

5.2.2 Overview of the datasets

We use three distinct datasets, which we match based on the *firm (business unit) identifier*. The matching procedure, resulting data selection and structuring of the data is described in detail in the Appendices. Here we overview and give the essential information about the data; descriptive statistics of the variables used in the final working dataset are presented in Table 5.1.

First, we make use of the "Company Surveillance Data" (referred to as VOV, its Danish acronym, throughout the rest of this paper), a 2001 survey on detailed workplace health & safety conditions and work environment practice, in a representative sample of Danish establishments within the private sector. The data covers information on subjective, general and specific, working environment status, and on various actions taken to address working environment problems. These answers are provided by a health and safety representative of the employees in each of the plants in the sample[10] and were col-

period of seven years, obligation for companies to assess their workplace conditions in the firm at least every three years, obligation for enterprises to seek for professional advice in workplace environment related matters etc.

[10]In Appendix B of this chapter we mention that we have two independent measures for each of the work environment indicators, given that both a health and safety representative from the side of the employees, and a health and safety representative from the managerial side, were asked to answer the work environment questionnaire. Analogous to Bloom and van Reenen (2006), we note that our two

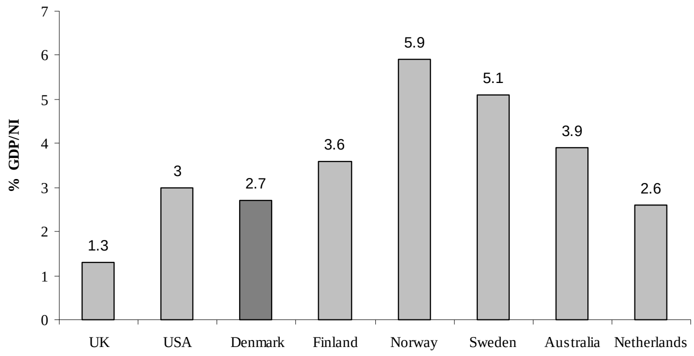

Figure 5.2: Estimates of aggregate economic cost of occupational injury and disease (%), by country

lected by persons specially trained for this type of surveys, from the National Institute for Occupational Health (AMI) in Copenhagen. Among the specific workplace dimensions covered we count problems related to "chemical loads", heavy lifting", "repetitive and strenuous work", "psycho-social" issues, "internal climate", "accidents and danger of accidents". Among the "work environment actions" undertaken, the representatives of the companies report on the firm's link to any formal occupational health and safety institute, whether the firm has a written working environment policy, whether general or specific work environment courses have been offered to the employees etc. A detailed discussion on the construction of the work environment indicators from the original questionnaire is presented in Appendix A. Although the VOV is collected at the establishment level, we are able to link it to the employer-employee and firm business account datasets only via the firm identifier, which means that we will be limiting our empirical analysis to firms with a single establishment[11]. The summary statistics table below contains therefore information on the sample of the mono-plant firms. In Appendix C of this chapter we show that the industry and geographic distribution of the firms with a single plant remains

independent measures for the specific workplace conditions have a fairly high correlation, which suggests that there isn't much bias in the individual answers. As explained in more detail in the Appendix, we choose to use for the empirical analysis the answers of the employees' health and safety representatives, given that there is somewhat more variation in these (the managers' representatives tend to rank work conditions as "good" or "very good" more often).

[11]Strictly speaking that is the case only for the analysis concerning the impact on firm productivity, where we need the REGNSKAB and IDA information; for the first, descriptive, part of the analysis we could use the sample of all plants from the VOV, if we do not want to say much about the aggregate employee and firm characteristics. We do want however to have the same mono-plant firms for comparability and hence we report results for the sample of mono-plant firms throughout the paper (results on the factors associated with a good working environment are essentially identical when using all the plants; estimates available from the authors upon request).

very similar to the initial dataset covering also the firms with more than one plant.

The second database used in this paper is the "Integrated Database for Labor Market Research" (IDA henceforth), constructed by Denmark Statistics from a variety of data registers used for the production of official EU and Denmark aggregate statistics. This data has been used and described in several previous studies, including Mortensen (2003), Bingley and Westergård-Nielsen (2003) or Buhai et al (2008). In a very brief depiction, IDA allows for matching of workers at establishments (local entities) and of establishments to firms (legal entities). It tracks every single work establishment and every single individual between 15 and 74 years old in Denmark. Apart from deaths and permanent migration, there is no attrition in the dataset. IDA is collected as of 1980 and includes detailed individual demographics such as gender, age, level of education, labor market state, experience, earnings, occupation, marital status etc.; other individual characteristics such as worker tenure can be reliably constructed, even if not present in the initial IDA. The labor market status of each person is recorded at the 30th of November each year. On the side of the employers, we have information on plant and firm employment size, region of firm location and industry category[12] and we can reliably construct a lower bound for the age of the firm (equal to the longest tenure among all of its employees). In this paper the information from IDA is used for constructing employee aggregate characteristics at the firm level, such as proportion of certain employee groups (i.e. proportion of females, unskilled, managers), mean and variance of education levels overall and per group, mean and variance wage in each firm, and the firm demographics indicators mentioned above[13].

Finally, we make use of a third dataset, on the firms' financial accounts. The statistics of business accounts (REGNSKAB henceforth), compiled by Denmark Statistics, cover construction and retail trade from 1994; the coverage was extended to manufacturing from 1995, to wholesale trade from 1998, and to the remaining part of the service industries from 1999 onwards. These statistics are aggregations of items of the annual accounts of business enterprises, notably items of the profit and loss account, the balance sheet and the statement of fixed assets. For the purpose of this paper we are specifically interested in the reported values for sales, capital stock and intermediate inputs (materials). There are several ways through which the statistics in REGNSKAB are gathered. The most thorough coverage is applied to firms that are selected for direct surveying; each year

[12] In our empirical analysis, we use the following broader industry indicators: 1. Agriculture & Mining; 2. Manufacturing; 3. Electricity, gas and water supply; 4. Construction; 5. Wholesale and retail trade; repairs 6. Hotels and restaurants; 7. Transport, storage and communications; 8. Financial intermediation; 9. Real estate, renting and business activities; 10. Public administration, defense and social security; 11. Education; Health and Social Work; 12. Other community, social and personal service activities

[13] The summary statistics table below presents only the "core variables" used in the final reported estimations in this paper.

these are all firms with more than 50 employees plus the firms with profits higher than a certain threshold, while smaller firms are included based on a rotation scheme. The firms are given the choice of either filling in a lengthy questionnaire or submitting their annual accounts plus detailed specifications. The questionnaire is modelled on the list of items set out in the Danish annual accounts legislation, so as to facilitate responding. The resulting data for the direct-surveyed firms are highly reliable. The other very reliable part of REGNSKAB is obtained from the tax forms submitted by firms, detailed enough for our purpose here. The rest of the data (typically the smaller firms - accounting for less than 20% of total turnover in the typical year) is obtained by stratified imputation based on employment size groups, with the method yielding results in large margins of error. For our paper we use therefore only firms directly surveyed and the firms where information has been obtained from their tax forms, implying again some data loss when linking to the other datasets. See Appendix D for an overview of the data loss due to the merger. For means and standard deviations of the variables of interest in the merged working dataset see the lower panel in the summary statistics table below.

Having overviewed the data, we stress that the *objective* variables in the two (independent) official datasets, IDA and REGNSKAB, are thus completely different in terms of source than the *subjective* workplace indicators contained in the VOV survey. This is a clear bonus vis-à-vis much of the earlier literature that used subjective measures of both dependent and independent variables, typically gathered from the very same respondents.

Table 5.1: Descriptive statistics VOV, IDA and REGNSKAB 2001

variable	definition	mean	s.d.	N
	VOV2001			
GENWE	work environment standard at the company is 5=VG,4=G, 3=NB, 2=B, 1=VB	3.86	.68	449
HLIFT	1 if problems related to heavy lifting have been solved, 0 otherwise	.76	.43	448
REPWO	1 if problems related to repetitive and strenous work have been solved, 0 otherwise	.78	.41	442
CHEM	1 if problems related to chemical loads have been solved, 0 otherwise	.88	.33	441
NOISE	1 if problems related to noise causing deafness have been solved, 0 otherwise	.77	.42	444
YOUNG	1 if problems related to young people's work have been solved, 0 otherwise	.94	.24	436
PSYCH	1 if problems related to psychological conditions have been solved, 0 otherwise	.78	.41	439
ICLIM	1 if problems related to internal climate have been solved, 0 otherwise	.71	.45	441
ACC	1 if problems related to accidents or danger of accidents have been solved, 0 otherwise	.80	.40	441
COURS	1 if courses with general work environment content have been held at the firm, 0 otherwise	.24	.43	426
ACTWE	1 if action plans have been drawn up to solve the work environment problems, 0 otherwise	.57	.49	437
PRIWE	1 if work environment problems have been prioritised to be solved, 0 otherwise	.68	.47	441
WRIT	1 if the firm has a Written Work Environment Policy, 0 otherwise	.32	.46	363
	IDA			
pfem	women as a proportion of all employees	.25	.26	572
pturn	employees with tenure less than two years as a proportion of all employees	.21	.22	572
punsk	unskilled as a proportion of all employees	.10	.19	572
pman	managers as a proportion of all employees	.078	.15	572
educ	average years of education among all employees	12.22	1.58	572
fsize	number of employees in the firm	49.37	96.09	572
agefirm	age of the firm	12.05	9.38	572
wage	mean wage in the firm	171.48	35.7	565
	REGNSKAB			
capital		15780.5	42698.6	465
sales		72639.3	190177	465
materials		54673.6	170577.6	465

5.3 Which are the factors associated with a good work environment?

In this section we focus on analysing the firm characteristics that may be correlated with the quality of the work health and safety environment in that firm, in other words we are investigating what differentiates firms with good workplace environment from the rest of the firms. To that aim we estimate different models that use alternative dependent variables as measure of the firm work environment quality. Our empirical methodology is analogous to Osterman (1994), who investigated the factors associated with the establishments' adoption of innovative work practices. Consider the following equation

$$WE_i = \alpha + \beta X_i + \gamma Z_i + \varepsilon_i \tag{5.1}$$

where WE_i represents the indicator of work environment health and safety quality for the ith firm, X_i is a vector of average firm and employees characteristics, Z_i is a vector of work environment actions that can improve workplace conditions and ε_i is an error term. Definitions and descriptive statistics for the variables used in our final specification can be found in Table 5.1 above.

We estimate logit models using both the general and all the specific work environment indicators. In all the estimations we transform the coefficients so that they have a direct interpretation, ie. we report the marginal change in probability of the specific work environment indicator being 1, given a one unit change in the independent variable[14].The first binary outcome model we estimate is contained in Table 5.2 (column1); the dependent variable is *GENWE*, an indicator taking value 1 if the "general work environment standard" at the company is "very good" or "good" and respectively 0 if it is "not bad", "poor" or "very poor"[15]. The only variable statistically significant at conventional significance levels is COURS[16], possibly suggesting that firms that held general courses with work environment content, with all the firm's employees, are more likely to increase the employees' awareness with respect to the work environment and thus ultimately obtain a better work environment compared to those that did not hold such courses. However, we

[14]The transformation is standard: $\frac{\delta p_i}{\delta x_{ij}}$ $=p_i(1-p_i)\beta_j$ with $p_i = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}}$; this expression is evaluated at the mean probability in the sample.

[15]We estimated also an ordered probit model with the dependent variable taking 5 values from "very good" to "very poor" and the results were qualitatively the same.

[16]Not shown in the estimates table for conserving space, the age or industry of the firm does not, surprinsingly, have any explanatory power in this general between-firm work environment differential either.

cannot give a causal interpretation to this result, aiming only to emphasize the statistical association in this exercise.

Columns (2) to (9) in Table 5.2 show estimates for a series of logits in which the dependent variables refer to specific work environment problems, with 1 if the specific condition "has been solved" and 0 otherwise. Most of the regressors take expected signs, but few are significant. The first covariate which is statistically significant is the log firm size: the larger the firm size the less likely are those firms "characterized by a good work environment", ie. having solved work environment related problems. The simple straightforward explanation for this result is that larger firms typically experience, in *absolute numbers*, more work environment related problems than smaller firms.[17].

The somewhat unexpected outcome is the importance the "proportion of managers" seems to have for several of the specific workplace health and safety indicators. In 3 equations (corresponding to HLIFT, REPWO and PSYCH) the coefficient on "*pman*" is positive and statistically significant, i.e. a higher proportion of managers in the firm is positively associated with better work environment in terms of heavy lifting, repetitive and strenuous work and psycho-social issues.

The variable COURS is again statistically significant for REPWO and NOISE and correlates positively with a good workplace environment, while the estimated coefficient on *pfem* is positive and significant for YOUNG suggesting that firms with a higher proportion of females in the workforce are less likely to face problems connected with young employees. Finally PRIWE, prioritizing work environment in the firm, is found positive and significant for the solution of problems connected to the internal climate[18].

An interesting remark is that many other aggregate firm characteristics (some of them not mentioned in the summary statistics table above for space reasons) do not have any power in explaining the between-firms workplace environment differential. What is perhaps most surprising is that such covariates like the proportion of "turnover employees"[19],

[17] This would be consistent with earlier literature where small and medium enterprises are the ones experiencing greater occupational safety and health problems *relative* to larger enterprises, see for instance Dorman (2000). This is for instance because often the improvement in workplace environment has substantial overhead costs and the smaller the firm, the smaller the revenue base over which these costs can be distributed; moreover, the formal work environment structures (eg. safety groups) and level of expertise in general is usually lower in smaller firms; finally, the market for SME's is usually more competitive, with finance more difficult to obtain, thus implying lower investment in general and particularly fewer expenditures on "non-essential" items.

[18] As in the case of the general work environment indicator above, the age of the firm is not found significant for any of the work environment specific dimensions. However, as expected, there are industry differences in this case. For instance the baseline category, agriculture, is clearly the worst in terms of "heavy load" problems, while chemical loads are worst for the manufacturing category etc.

[19] As defined in Table 1, in our data *pturn* represents the employees with tenure less than two years, as a proportion of all employees (hence, employees who just entered the firm and are observed for the first time in the data).

"having a written working environment policy", "mean education of managers" , "mean experience of the managers" (both these latter ones potentially proxying manager ability), "mean tenure in the firm", "variance in the age composition", "average firm tenure" or experience, are not statistically significant [20].

The findings from our specifications above suggest that there are only a couple of robust variables positively associated with most specific measures of good workplace environment. Namely, these are the higher proportion of managers and respectively, offering courses with work environment content. To less extent, the proportion of females within the firm and prioritizing work environment practice in the firm also seem to explain across-firm differences in some of the work environment dimensions. If we are willing to speculate somewhat, our conclusions herein could be interpreted in the sense that the higher proportion of managers being positively associated with better workplace conditions indicates the beneficial effect of managerial involvement in workplace environment related issues and, analogously, that raising employee awareness by means of work environment related courses can also raise workplace conditions. In fact, these two factors could well be complementary within a firm, as supported for instance by studies such as Kato and Morishima (2002), who provide evidence on the association between top-level management and shop-floor employee participation in workplace organization decisions.

[20] Results for all alternative models using these variables are available on request from the authors.

Table 5.2: Logit estimates of work environment on firm characteristics in 2001, marginal effects

	GENWE	HLIFT	REPWO	CHEM	NOISE	YOUNG	PSYCH	ICLIM	ACC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
pfem	.083	.041	.140	.035	-.025	.066*	.023	-.195	.225
	(.119)	(.131)	(.118)	(.038)	(.076)	(.036)	(.121)	(.137)	(.165)
punsk	-.0007	.205	-.184	-.016	.163	-.051	.026	.195	-.118
	(.148)	(.179)	(.165)	(.057)	(.128)	(.065)	(.134)	(.199)	(.139)
pturn	-.020	-.030	.137	.004	-.044	-.023	-.207	-.219	.003
	(.150)	(.125)	(.146)	(.051)	(.091)	(.032)	(.134)	(.166)	(.141)
pman	.049	.640**	.654**	.074	.191	.129	.676**	.139	.309
	(.204)	(.310)	(.302)	(.079)	(.141)	(.101)	(.294)	(.259)	(.235)
educ	.014	-.008	.010	-.015	.012	-.001	-.004	-.027	.001
	(.019)	(.021)	(.022)	(.012)	(.012)	(.007)	(.023)	(.026)	(.026)
log fsize	-.019	-.110***	-.076***	.001	-.041***	-.008	-.091***	-.091***	-.078***
	(.023)	(.026)	(.024)	(.007)	(.015)	(.006)	(.024)	(.026)	(.026)
COURS	.194***	-.069	.103*	.020	.082**	-.025	-.023	-.023	-.037
	(.056)	(.060)	(.059)	(.015)	(.033)	(.016)	(.055)	(.062)	(.052)
WRIT	.059	.080	.017	-.016	.040	-.002	-.026	-.031	.051
	(.062)	(.063)	(.067)	(.015)	(.036)	(.015)	(.062)	(.063)	(.058)
ACTWE	.089	.053	.090	-.010	-.022	-.002	-.016	-.129	.061
	(.067)	(.070)	(.079)	(.017)	(.047)	(.017)	(.074)	(.085)	(.064)
PRIWE	.104	-.045	-.073	-.008	.033	-.005	.048	.176*	-.106
	(.070)	(.080)	(.083)	(.023)	(.053)	(.019)	(.077)	(.096)	(.074)
Nobs	305	279	280	230	279	175	297	302	264
Log-lik	-147.06	-129.56	-125.14	-66.86	-129.48	-42.57	-140.79	-152.98	-117.36

Significance levels: *** 1%,**5%, *10%; White heteroskedastic-consistent standard errors in parantheses. Estimations also include a constant term, regional and industry dummies and dummies for firm age categories, ie. age 0-5, 5-10, 10-15, 15-20, with the baseline 20+

5.4 Impact of work environment on firm performance

5.4.1 Impact on firm productivity

In the second part of the paper we are first interested in the determinants of the firm's total factor productivity, focusing on the role of the workplace's health and safety environment. To that aim, we will be estimating standard Cobb-Douglas production functions, augmented with the firm specific workplace environment indicators used as dependent variables in the binary outcome regressions from the previous section, and with employee aggregate characteristics. Our analysis largely traces the two-step empirical strategy by Black and Lynch (2001), technique that has also been recently used in a related context by Zwick (2004). Namely, although VOV is cross-sectional, we can make use of the information compiled from IDA and REGNSKAB for previous years as well, and hence are able to estimate three distinct specifications for the production function.

The simplest specification is using only the cross-sectional sample with all variables, i.e. estimating the following OLS regression:

$$\ln(Y/L)_i = c + \alpha \ln(K/L)_i + \beta \ln(M/L) + \delta' X_i + \gamma' Z_i + \varepsilon_i \tag{5.2}$$

with c a constant term, Y/L sales per firm size, K/L capital per firm size, M/L intermediate inputs (materials) per firm size, vector X containing the firm specific characteristics of employees and vector Z containing our establishment specific workplace practices [21]. We use the stock value of capital K and intermediate materials M reported in the REGNSKAB database[22]. The results of the estimation above are reported in column(1) of Table 5.3. All the OLS estimations control also for location, industry and age of the firms.

Since our cross-sectional estimates from (5.2) may be subject to endogeneity due to unobserved heterogeneity in the firm characteristics that above is all captured by the error term ε_i, we exploit further the fact that we observe the IDA and REGNSKAB datasets of our firm aggregate variables over time, in order to eliminate any unobserved *time-invariant* firm fixed effects, and use the residual from the first stage, averaged over time (ie. the time-invariant component of the residual), as dependent variable in a second stage OLS regression on the 2001 cross-section of work environment indicators[23]. The

[21] We verify that the constant returns to scale restriction is not rejected in our data. Unlike Black and Lynch(2001), we cannot clearly distinguish between "production" and "non-production" workers in our data, hence we will use the general specification using all the firm's labour force.

[22] K is computed by adding the intangible and tangible fixed assets; M is calculated as sales minus value added, using the value added formula provided provided by Denmark Statistics.

[23] Just like in Black and Lynch (2001), in the first stage we have the option of using all the available

empirical specification in this case is given by:

$$\begin{aligned}
\widetilde{\ln(\frac{Y}{L})}_{it} &= a\widetilde{\ln(\frac{K}{L})}_{it} + b\widetilde{\ln(\frac{M}{L})}_{it} + c\widetilde{X}_{it} + \widetilde{\nu}_{it} \text{ (step1)} \quad (5.3)\\
R_i &= d + e'Z_i + \xi_i \text{ (step 2)}\\
\text{where } R_i \text{ is the (time) average of } R_{it} &\equiv \widetilde{\ln(Y/L)}_{it} - \widehat{a}\widetilde{\ln(K/L)}_{it} - \widehat{b}\ln\widetilde{(M/L)}_{it} - \widehat{c}\widetilde{X}_{it}
\end{aligned}$$

where the upper tilde means that we use deviations from the means over time[24]. Note that we differ already slightly from Black and Lynch (2001), in that we also observe the firm aggregate employee characteristics over time, and thus can use them as well in the first stage regression. The values for sales, capital and materials were deflated using the net price index provided by Denmark Statistics, with a base year of 2000. In the reported results we use $t = \overline{1998, 2001}$, since this is a very likely period over which the work environment indicators are not expected to vary[25]. However, varying the time period by including also earlier periods (earliest available is 1994, but that includes very few establishments also observed in 2001) or using less lags does not affect the qualitative interpretations of the results. The results of this second empirical strategy are presented in column (2) of Table 5.3.

Although the specification from (5.3) above would take care of any time-invariant firm effects that could be correlated with the choice of inputs in the first stage, the typical simultaneity problem in choosing the production inputs or the measurement error in the explanatory variables (capital and materials) has still not been dealt with. The pitfall in production function estimation, known since Marschak and Andrews (1944), is the

observations (including observations for establishments with missing information on certain work environment indicators in the 2001 cross-section) or just the observations from the establishments used in the second stage. Since results are identical with either alternative (less so the magnitude of the standard errors in the first stage regression, but they do not affect the statistical significance interpretation of the point estimates for any of our variables), we report the 1st stage results for the larger sample.

[24] We assume that ν_{it} is a disturbance with 0 mean, so that taking deviations from the average over time eliminates or considerably reduces its contribution to the residual.

[25] A provision in the Danish Work Environment Act states that workplace assessments shall be undertaken "at least every 3 years", which suggests that 1998-2001 is a likely period on which to expect workplace indicators not to change much. This expectation is enforced also by the fact that another question in the VOV questionnaire, asking about the last time a work environment assessment was implemented and what types of problems were found at that time, suggests that 60 to 80% (depending on the specific work environment indicators) of the observed work environment indicators do not change since the last assessment (there are many missing values however). Moreover, most previous work environment assessments, if the question on the timing is answered (many missing values however also here), are indeed reported to have been implemented in the interval 1998 to 2001. Note that the length of this time period is shorter than in the case of work reorganization measures as analyzed in Black and Lynch (2001, 2004). This is not unusual, given the faster expected impact of changes in workplace environment conditions than that of crucial changes in the organization of the entire production process, for instance.

endogeneity of input choices in the production function, given their likely correlation to unobserved productivity shocks, c.f. Griliches and Mairesse (1998). To address that, analogous to Black and Lynch (2001), we exploit the fact that we can observe all variables (except the ones from the VOV dataset) over time, to apply a system-GMM estimation à la Arrelano and Bover (1995) and Blundell and Bond (1998, 2000) in the first stage, and to subsequently use the averaged residuals over time from this first stage as dependent variable in a second stage, as an OLS on the vector Z, containing the work environment indicators. This approach involves estimating the 1st stage from expression (5.3), without the upper tilde on the variables, by using appropriately *lagged values of both levels and changes* in capital, material, labour and output, as *instruments for levels* of capital, material and labour. Furthermore, as a plus relative to Black and Lynch (2001), given that the proportion of managers in a firm was strongly associated with a firm having a good workplace environment for most workplace indicators (see the previous section), we are also instrumenting with lagged levels and changes of that variable; this proxies for the time-varying "managerial ability" that might still remain correlated with the work environment indicators in the final stage of the estimation procedure. The estimates of this latest strategy are presented in the third column of Table 5.3, where again we use time lags down to 1998, as in the fixed-effects strategy from the previous column. We first check that the conditions for applying the system-GMM are in place: the validity of the instruments and respectively, the assumption of no serial correlation in the levels error term ν_{it}. According to the Sargan-Hansen test for overidentifying restrictions, we do not reject the validity of our instruments at conventional statistical levels. We also do not reject the null hypothesis of no serial correlation in ν_{it}; since the reported LM tests are performed for the differenced residuals $\Delta\nu_{it}$, c.f. Arrelano and Bond (1991), we are interested in confirming the absence of the second order serial correlation, whereas the negative first-order serial correlation is consistent with our specification, see also Dearden et al (2006).

What can be learnt from the estimations in Table 5.3? Firstly, whether we instrument the proportion of managers GMM-style (the reported estimate in the table is for this case) or we do not, does not influence at all the results; hence, time-varying managerial ability (at least as proxied by proportion of managers over time) does not appear to matter in this production function estimation. Secondly, a number of results are completely consistent with the findings in Black and Lynch (2001). Thus, we notice that our point estimate for K/L increases from the 1st (simple OLS) to the 3rd 2-stage (OLS+ system GMM) estimation strategy, as expected, suggesting that indeed the latter empirical specification accounts to some extent for the fact that in the previous two strategies the

Table 5.3: Augmented production functions

	OLS 2001	2-stage FE+OLS	2-stage GMM+OLS
	(1)	(2)	(3)
		1^{st} stage	
K/L	.034*	.048***	.060**
	(.017)	(.011)	(.027)
M/L	.671***	.751***	.745***
	(.026)	(.022)	(.061)
pfem	.002	-.053	-.053
	(.106)	(.053)	(.053)
punsk	-.262**	-.022	-.013
	(.111)	(.033)	(.036)
pturn	-.138	-.082***	-.096***
	(.130)	(.021)	(.035)
pman	.329	.017	.127
	(.217)	(.075)	(.187)
educ	.002	-.006	.003
	(.016)	(.006)	(.008)
Nobs 1^{st}stage		1627	1627
Sargan			$\chi^2(15)$=19.40 (p-value=0.20)
LM 1^{st} order serial corr			z=-3.65 (p-value=0.00)
LM 2^{nd} order serial corr			z=-0.30 (p-value=0.77)
		2^{nd} stage	
WRIT	.021	.018	.011
	(.031)	(.030)	(.029)
COURS	.044	.043	.040
	(.035)	(.034)	(.034)
ACTWE	.004	-.0006	.022
	(.047)	(.048)	(.046)
PRIWE	-.030	-.028	-.029
	(.046)	(.047)	(.046)
HLIFT	-.021	-.035	-.041
	(.044)	(.044)	(.044)
REPWO	.070	.094**	.092**
	(.045)	(.042)	(.042)
CHEM	.074	.058	.059
	(.073)	(.063)	(.063)
NOISE	-.008	.010	.006
	(.035)	(.031)	(.030)
YOUNG	-.022	-.043	-.043
	(.047)	(.041)	(.040)
PSYCH	-.025	-.013	-.012
	(.036)	(.037)	(.035)
ICLIM	.041	.074**	.080**
	(.037)	(.031)	(.031)
ACC	.011	-.008	-.015
	(.036)	(.031)	(.030)
R^2	0.920	0.225	0.242
Nobs	215	215	215

Significance levels: *** 1%,**5%, *10%; White heteroskedastic-consistent standard errors in parantheses. Estimations also include a constant term, regional, industry indicators and dummies for age categories of the firm. For the 1st stage FE and GMM regressions we also control for interaction between year and industry dummies. Sargan is a χ^2 test of overindentifying restrictions; LM is a Lagrange Multiplier test of 1^{st} and respectively 2^{nd} order serial correlation in Δv_{it}, distributed N[1,0] under the null; p-values for the significance test of the null hypotheses are reported in brackets, after the test coefficients estimates were more tainted by measurement error[26]. Next, we also find that only the "proportion of turnover employees" is statistically significant and of the expected sign, among our common aggregate worker characteristics in the production function estimates. Also consistent with Black and Lynch (2001), most of the results concerning the effect of aggregate employee characteristics are qualitatively and quantitatively robust over both the FE and system-GMM specifications in columns 2 and 3. Finally, what can we say in terms of the impact of the workplace health and safety environment, our main concern? In both the fixed effects and the system-GMM specifications we find that the only work environment actions that matter are having solved "internal climate problems" and respectively, having solved problems concerning the "repetitive and strenuous work", both having rather large marginal contributions relative to the other production inputs. This suggests that these two physical dimensions of the work safety and health environment tend to be critical for the firm's total factor productivity, while psycho-social dimensions as well as other work environment criteria such as general work environment status, do not seem to contribute at all to enhancing firm performance.

5.4.2 Impact on mean wages

The other indicator for "firm performance" that we are going to look at in this study is the firm's mean wage, a classical proxy for the employees' welfare. This is obtained from IDA, averaging over the hourly wages of all workers in the firm[27].

We are interested in the extent to which differentials in mean wages offered by the firms are explained by work environment conditions and by other aggregate employee characteristics. Comparing the mean wages of firms that implement good work health and safety practice to those that do not directly by nonparametric propensity score matching– previously used in the literature in similar contexts, e.g., Janod and Saint-Martin (2004)– is not feasible here given the rather low sample sizes of our working datasets. Hence, we will implement two simple strategies using log mean wage as dependent variable, following the methodology used in the previous subsection, on firm productivity. The first method is to use OLS in the cross-sectional 2001 sample, while the second consists in

[26] Our point estimate for K/L is still on the low end of what is found in the literature, even when using the sys-GMM. Using a back-of-the-envelope computation, our estimates would suggest that capital accounts for rougly 1/4 of value added (sales minus intermediary material costs) and labour for the rest. There are however also other papers that have found even lower capital intensities in such augmented production frameworks, using the same system-GMM technique, see for instance Zwick (2004).

[27] We take care of the outliers in wages by trimming the top 1 percentile of the cross-sectional wage distribution for that specific year and truncating all reported wages below the legal minimum wage in that year. For the empirical specification where we use different time periods, we deflate wages with the consumer price index using 2000 as base year.

exploiting the fact that we observe variables obtained from IDA over time, and hence we can use that information to develop a 2-stage estimation analogue to the second estimation strategy from the previous subsection, where in the first stage we recover a firm fixed component of the residual and we use it as dependent variable in the second stage, with the workplace environment indicators as independent variables. The second strategy takes care of any unobserved time-invariant firm heterogeneity that might be correlated with the firm specific characteristics. The above can be written as

$$\text{OLS: } \ln(Y)_i = c + \acute{\alpha} X_i + \beta' Z_i + \varepsilon_i \tag{5.4}$$

$$\begin{aligned} &\text{2-stage, FE+OLS:} && \\ \widetilde{\ln(Y)}_{it} &= \acute{a}\widetilde{X}_{it} + \widetilde{\nu}_{it} \;\; \text{(stage 1)} \\ R_i &\equiv d + b' Z_i + \xi_i \;\; \text{(stage 2)} \\ \text{with } R_i \text{ the (time) average of } R_{it} &\equiv \widetilde{\ln(Y)}_{it} - \widehat{\acute{a}}\widetilde{X}_{it} \end{aligned} \tag{5.5}$$

where c and d are constant terms, vector X collects the firm specific characteristics, vector Z contains the work environment proxies, v is a time-invariant firm effect and ε, ν and ξ are error terms. Y is mean wage. $\widehat{a}$ is the estimated value of a from the first stage. The upper tilde indicates that we take the deviations from the means over time[28]. All OLS estimations control for regional, industry and age of the firm indicators.

The estimates for logwages as dependent variable are in Table 5.4; the first column contains estimates of the OLS, the second contains estimates of the two-stage FE+OLS estimation.

What is the interpretation of the log wage regression estimates? First, there are some differences between the cross-sectional estimates and the estimates using the 2-stage strategy (the effect of the aggregate employee characteristics is identified from variations over time in this latter case, since they are included in the first stage). Thus, "proportion of managers" is significant in first column, but ceases to be significant when we use the 2 step FE+OLS technique from the 2nd column; there is a similar case with "having problems related to young employees (YOUNG). Since the second stage takes into account possible unobserved time-invariant heterogeneity in the employee characteristics that could be correlated with the workplace environment indicators, we prefer the 2-step specification. Other conclusions are carrying over from the 1st to the 2nd column and confirm pervasive

[28] We use 1998-2001 as the time period in the reported estimates, although the results are virtually identical when we vary it, including less or more lags (earliest possible being 1994).

Table 5.4: Mean logwages and work environment

	OLS 2001	2-stage FE+OLS
	(1)	(2)
		1^{st} stage
pfem	-.212***	-.159***
	(.051)	(.048)
punsk	.102	.015
	(.057)	(.035)
pturn	.022	-.0004
	(.054)	(.016)
pman	.335***	.012
	(.092)	(.053)
educ	.051***	.038***
	(.009)	(.006)
Nobs 1^{st}stage		2095
		2^{nd} stage
WRIT	.020	.026
	(.023)	(.023)
COURS	.018	.020
	(.020)	(.020)
ACTWE	-.009	.018
	(.028)	(.023)
PRIWE	.022	.015
	(.033)	(.028)
HLIFT	.029	.015
	(.022)	(.021)
REPWO	-.035	-.011
	(.022)	(.021)
CHEM	-.004	.026
	(.034)	(.027)
NOISE	-.011	-.015
	(.022)	(.022)
YOUNG	.054*	.024
	(.032)	(.030)
PSYCH	.019	.034
	(.023)	(.022)
ICLIM	-.020	-.040*
	(.023)	(.024)
ACC	.030	.019
	(.026)	(.025)
R^2	0.491	0.323
Nobs	295	295

Significance levels: *** 1%,**5%, *10%; White heteroskedastic-consistent standard errors in parantheses. Estimations also include a constant term, regional, industry indicators and dummies for age categories of the firm. For the 1st stage FE regression we also control for interaction between year and industry dummies.

results throughout the empirical literature: a higher proportion of female employees is strongly associated with a lower mean wage at the firm, while a higher mean employee education translates in higher firm mean wages. Could there be any compensating differentials story to be told? In the cross-section OLS estimation none of the work environment indicators turns out to matter, except YOUNG (with a positive sign), but that becomes statistically not different from 0 in the second column. However, in our preferred 2nd column of estimates, having solved "internal climate" conditions, is associated with a lower wage, which might indicate the fact that bad internal climate is compensated for by higher mean wages. Quite surprising is that none of the other work environment indicators or other employee aggregate characteristics appears to explain the mean wage differentials across firms.

5.5 Summary and discussion

This is the first paper to investigate which are the firm characteristics associated with a good workplace health and safety environment and what is the impact of such good work environment practice on firm performance, both in terms of total factor productivity and firm mean wage. We have merged Danish data from three independent sources to investigate: a. which aggregate employee characteristics can explain the between-firm differential in workplace environment and b. what is the impact of improving workplace conditions on firm productivity and firm wages. Our findings suggest, on the one hand, that few factors are associated with a good work environment practice, but that those found relevant are important across several work environment indicators. The main factors are the proportion of managers and respectively, courses with work environment content offered to all the employees. The first factor might suggest that high management involvement is important, while the second might indicate the role the employees' awareness plays, in enhancing workplace conditions. Other factors that seem to matter less are the proportion of female employees and prioritizing work environment practice at the firm. On the other hand, we have found that the explanatory power of work environment related practice in explaining between-firm wage differentials is rather low. Once we control for industry, regional and firm age effects, the only work environment dimension accounting for a compensating wage differential story is the internal climate at the workplace. The conclusion regarding the importance of this physical dimension of the workplace environment is consistent also in the light of the firm productivity estimates. According to the results from the production function estimations, the work environment related factors that contribute to enhancement of firm productivity are having solved problems related to "internal climate" and respectively, to "repetitive and strenuous work", both with relatively large marginal contributions to enhancing firm productivity.

As Black and Lynch (2001), we are aware that neither of our 2-step methods cannot fully account for possible endogeneity of the work environment indicators in the production function: some time-varying unobserved heterogeneity correlated both with firm profits and work environment indicators could, theoretically, still bias our final estimates. However, in practice, it is not easy to think of a clear source for such further omitted variable bias: in addition to the careful methodology borrowed from Black and Lynch, we have also fully exploited the fact that in our data we observe all aggregate employee characteristics over time. In particular we have been able to instrument the current proportion of managers in our system-GMM procedure with its changes and lagged levels, which could be thought of as proxying time-varying managerial ability of the firm.

It will be interesting to see similar future studies using different datasets and comparing their findings to the ones in this paper. In particular, ideally one would like to be able to use longitudinal observations also on firm workplace health and safety conditions, next to observing all other firm characteristics over time. Given the enormous aggregate costs of job-related accidents and illnesses in all developed nations, it is obvious that corporations, trade unions and policy agents should all be very interested in the outcomes of such research, hence we do not expect to remain the only paper in this area for long.

Appendices Chapter 5: Data selection and structuring

5.A Construction variables VOV

The main dataset in the merging procedure is VOV. Herein we describe the construction of the variables in this dataset.

The key variables of interest are working conditions indicators constructed from the questionnaire; these indicators cover aspects such as physical, psychological strain and danger of accidents.

A set of dummies regarding specific work environment problems is created, that take value 1 if the firm indicates that the "majority" of problems have been solved and value 0 if "few" or " none" problems have been solved[29]. These variables are developed from the question " To what extent problems related to heavy lifting ($HLIFT$)/ repetitive strenuous work ($REPWO$)/ chemical loads($CHEM$)/ noise causing deafness ($NOISE$)/ problems in connection with children and young people´s work ($YOUNG$)/ psychological conditions[30]($PSYCH$)/ internal climate problems and accidents ($ICLIM$)/ accident or danger of accident (ACC), have been solved ?". On average, about 75% of firms report that the majority of the specific work environment problems have been solved.[31]

A subjective "general work environment status" indicator GENWE is constructed from the question "What do you consider the work environment related standard to be at the company? very good/good/not bad/poor/very poor", and takes value 1 if the general work environment standard at the company is very good or good and value 0 if it is not bad, poor or very poor.

Another set of dummy variables describes various actions undertaken in connection with the work environment, such as WRIT, which is derived from the question "Does the company have a written work environment policy?yes/no/don´t know)"; COURS,

[29] We note here that we do not know precisely *when* these problems have been actually solved, hence we cannot perform, e.g., an analysis of changes in firm performance on changes in these indicators, since we do not know which lagged time period to use in order to compute changes in firm performance (or other firm characteristics). What we know from another question in this survey is that the last workplace environment assessment took place within the last three years for most firms in the sample (there is also a Danish organic law that states that these assessments should be done at least every 3 years) and that at this last assessment some of these problems were reported not to have been solved (20 to 40% depending on the precise workplace indicator); unfortunately we have too many missing observations in order for an empirical analysis using changes in the workplace indicators from the last assessment (whenever that was) to be feasible.

[30] From conversation with the people who designed the questionnaire we know that these include issues such as pressure of time, lack of influence, work times, working alone, perceived violent/uncooperative environment etc.

[31] For mono-plant firms only we get the same proportion, compare e.g. Table 5.1, in the text.

"Has the company\workplace held courses, project days, seminars or similar events for its employees where the work environment has to a greater or lesser extent been included as a subject? yes/number of events in the last year/no/don´t know"; ACTWE "Have you drawn up action plans to solve the work environment problems?yes/no/don´t know" and PRIWE, "Have you prioritised the work environment problems that are to be solved?yes/no/don´t know".

5.B Employee representative vs. employer representative in VOV

The VOV 2001 questionnaire is asked to both one safety group representative of the employees ("type 1") and one safety group representative of the employer ("type 2"), for each establishment, so that the initial data contains two observations for each establishment surveyed. The first selection step is that we only keep the answer of the employees' safety representatives and we do not use the second measure, though note that they are fairly highly correlated for the specific work environment measures (the correlation coefficient is between 0.35 and 0.70 for each of these specific safety and health measures, with an average across all of them slightly higher than 0.50). Our decision is mainly motivated by the fact that the variation in answers of type 1 is somewhat higher than the ones in type 2, with the latter tending to cluster around "very good" or "good" for most questions. Since the questionnaire related to health and safety assessment of the workplace, we believe the workers' answers to be the ones more reliable[32]. To illustrate the difference in the variance between the two types with one (extreme) example, consider the answer to the general question concerning the work environment related standard (the correlation between the two measures for this general work environment indicator is only 0.17). Table 5.5 present the answers of both "types" to the question: "What do you consider the work environment related standard to be at the company?", for observations where both types's answers are nonmissing. We define an ordered variable defining the general work environment (*GENWE*), taking values that range from 1=very good to 5=very poor.

From Table 5.5 it appears clear that type 1 answers have more variance than type 2 answers[33], although the difference is lower for all of the specific work environment indica-

[32] One rationale for that is the fact that previous research has clearly documented that employee attitudes at the workplace can have significant impact on economic outcomes at those firms, see for instance Bartel et al (2003); hence, we would precisely like to use the answers of the employees' safety and health representative for our investigation.

[33] The discrepancy remains the same if we consider only the mono-plant firms, the ones used in the empirical analysis.

Table 5.5: Differences between types, all plants

GENWE	Type=1		Type=2		Total	
	N	%	N	%	N	%
very good	156	13.15	362	30.57	518	21.9
good	707	59.17	693	58.53	1400	59.1
not bad	286	24.16	123	10.39	409	17.3
poor	30	2.53	5	.42	35	1.5
very poor	5	.42	1	.08	6	.2
Total	1184		1184		2368	

tors. In fact, performing all our estimations with the answers of type 2 we get identical qualitative results, with the exception that in some cases the statistical significance is lost if using the employer representative's answers[34].

5.C Mono-plant firms vs. multi-plant firms in VOV

Given that we have to match the datasets on the firm identifier, we select only firms that have a single establishment (plant) for the rest of the analysis. How representative does this sample remain of the private Danish sector in terms of geographical and industry distribution? The two tables below show respectively the distribution by industries, Table 5.6., and the distribution by regions, Table 5.7., for both the initial sample of all plants and the working sample of mono-plant firms. We notice that the mono-plant firms keep largely the same geographical distribution as the plants in the initial sample and that the only considerable changes are in the case of two industries: for "real estate" where the proportion of plants decreases from 4.3% of the total sample, initially, to 2.4%, in the working sample, and especially for the private firms operating in the" public administration, defense and compulsory social security" category, where the plant percentage decreases from 5% in the initial sample to 0.7% in the working sample of mono-plants.

[34]We also note here that an empirical strategy in which one would instrument one of the measures with the other one, is not directly feasible given that we deal with ordinal (and mostly binary) indicators here, as well known in the econometrics literature. Moreover this strategy would be dubious as well, in the light of our goal: if anything, it is likely that eventually both these measures would be correlated with some unobserved time-varying measure of managerial ability, and thus, with firm performance and hence the validity of the instrument is not met.

Table 5.6: Distribution by industries

	All-plants		Mono-plant firms	
	N	%	N	%
Agriculture, fishing, mining and quarrying	33	3.6	27	4.7
Manufacturing	546	59.7	357	62.4
Electricity, gas and water supply	1	0.1	1	0.2
Construction	59	6.5	47	8.2
Wholesale and retail trade	68	7.4	45	7.9
Hotels and restaurant	5	0.5	4	0.7
Transport, post and communication	45	4.9	32	5.6
Financial intermediation	17	1.9	6	1
Real estate, renting and business activities	39	4.3	14	2.4
Public administration, defense and social security	46	5	4	0.7
Education	32	3.5	20	3.5
Health and social work	12	1.3	6	1
Other community, social and personal service activities	11	1.2	9	1.6
Total	914		572	

Table 5.7: Distribution by regions

	All plants		Mono-plant firms	
	N	%	N	%
Copenhagen	197	21.6	103	18
Roskilde	24	2.6	19	3.3
Vestsjaeland	54	5.9	36	6.3
Storstroem	27	3	22	3.8
Fyn and Bornholms	105	11.5	53	9.3
Soenderjylland	67	7.3	43	7.5
Ribe	56	6.1	32	5.6
Vejle	73	8	50	8.7
Ringkoebing	43	4.7	28	4.9
Aarhus	69	7.5	48	8.4
Viborg	97	10.6	64	11.2
Nordjylland	102	11.2	74	13
Total	914		572	

5.D Data loss in merging VOV-IDA-REGNSKAB

We face some unavoidable sample reduction during the merging procedure, which we briefly describe below:

- We start with 1962 establishments sampled in VOV 2001 (we have two observations for each of these establishments, corresponding to type 1 and type 2, as explained earlier in this Appendix).
- We need to find the firm identifier for most of the initial establishments, since these were often sampled in the dataset only by their name and that string was sometimes entered only partially in the database etc. This was done (by a very tedious manual work performed by very patient student research assistants) using an auxiliary business statistics dataset (known as KØB), matching names to firm identifiers. We were not able to find the firm identifier for 490 of the initial establishments.
- We need to use only mono-plant firms in merging to IDA and REGNSKAB, since we do not have establishment identifiers in VOV to match directly with establishments in IDA and since in REGNSKAB we have of only business account statistics at the business unit, that is the firm level. That leaves us with a sample of 572 firms in the merged VOV-IDA dataset and 465 firms in the merged VOV-IDA-REGNSKAB dataset. We have less firms in REGNSKAB given the sampling procedure in the construction of that dataset and its reliability only for part of the firms, see also the REGNSKAB overview in the data description part of this paper.
- For the production function estimation we use all the available observations in VOV-IDA-REGNSKAB, while for the impact on mean wages, we use all the available observations in VOV-IDA. In the empirical analyses we end up de facto with even smaller sample sizes, given that many of our variables used in the estimation have missing observations.

Chapter 6

Concluding remarks

This book is a compilation of four autonomous essays, spanning several research areas within the– largely defined– field of *labour economics*. Its ambition has been to provide sound and original research perspectives on each of the topics tackled, some of them classics in the economics discipline. I believe this thesis has more than fulfilled that aim: it has not only managed to contribute with innovative methodological approaches, both theoretical and empirical, but in several cases it has obtained some brand-new results, often counterintuitive in light of the existent economics know-how. Moreover, some of the findings herein can easily be imagined as having potential implications for both corporate and policy practice, though I also think that one should never tire of stressing that the main and immediate intention has been to contribute to the academic economic literature, and to see future studies following up; too often academic economists seem too eager to promote fresh research beyond the academe, before this work has gotten the slightest chance to be well digested by scholars.

Chapter 2 provides a new way of analysing tenure profiles in wages, by modelling simultaneously the evolution of wages and the distribution of tenures. We develop a theoretical model based on specific initial investments at the job and efficient bargaining on the job surplus, where both log outside wage and log wage in the current job follow a random walk, as verified empirically. This setting allows us to apply real option theory. We can derive the efficient separation rule between the worker and the firm. Our model fits the observed distribution of job tenures very well. Given that we observe outside worker wages only at job start and job separation, our empirical analysis of within job wage growth is based on expected wage growth conditional on the outside wages at both dates. We test our model on an extract from the Panel Study of Income Dynamics and obtain a surprisingly good fit, although we point out that wages exhibit downward rigidity, suggesting that a more suitable model should take that feature into account; this is an extension we leave for future work. One of our main conclusions is that an empirical analysis not taking into account the selection on the outside option of the workers would indeed provide high estimated tenure profiles in wages, in the order of typical OLS cross-sectional estimates. However, we show that almost all that return takes the form of a declining outside productivity, instead of a rising inside productivity, and by excluding that part of the tenure profile our estimates would instead be on the very low end of the spectrum.

In Chapter 3 we document two empirical regularities, using exhaustive longitudinal match employer-employee data for Denmark and Portugal. First, workers who are hired last, are the first to leave the firm: Last-In-First-Out applies. Second, the workers' wages unambiguously rise with seniority (= a worker's tenure *relative* to the tenure of her colleagues), the traditional return to the time spent in the job being just a measurement error due to misspecification of earlier models. We explain these regularities by developing a dynamic model of the firm with stochastic product demand and hiring cost, i.e. irreversible one-time specific investment for our purpose. There is wage bargaining between a worker and its firm. Job separations (quits or layoffs) obey the LIFO rule and bargaining is efficient, implying a zero surplus at the moment of separation. The LIFO rule provides a stronger bargaining position for senior workers, leading to a return to seniority in wages. Efficiency in hiring requires the workers' bargaining power to be in line with their share in the cost of specific investment. The LIFO rule can thus be interpreted as a way to protect their property right on the specific investment. We also discuss the effects of firing costs, Employment Protection Legislation and risk aversion. One very interesting corollary of our model is that a return to seniority implies that a worker is to an extent shareholder in her own firm. While we have established the existence of a return to seniority for Denmark and Portugal, an obvious caveat is that we cannot say whether such a return exists in other countries, in particular in the United States. However, there are strong indications that this should be the case: the large return to job tenure in the United States, as compared to Denmark and Portugal, strongly suggests so. Moreover the economic mechanisms for having a LIFO layoff rule exist everywhere, and legal institutions specific to Denmark or Portugal but not to other places might simply be a formalisation of rules of conduct and implicit contracts that would have emerged anyway.

In Chapter 4 we develop a simple social network model of occupational segregation between social groups divided along gender, race or ethnical origin. Jobs are obtained through a network of contacts formed stochastically, but with intra-group inbreeding bias, after career decisions had been made. We establish that even with a minimal amount of homophily within each social subgroup, stable occupational segregation equilibria will arise. Moreover, if the wage differential across the occupations is not too large, complete segregation will always be sustainable. If the wage differential is large, complete segregation cannot be sustained, but a partial segregation equilibrium in which one of the group fully specializes in one education while the other group mixes over the career tracks, is sustainable. Furthermore, our model is able to explain sustained unemployment and wage differences between the social groups. We also consider implications of our model from a social planner's point of view. Quite surprisingly, we find that socially optimal policies involve segregation in both the first and second best cases. In the first best optimum, we find that segregation is the socially preferred outcome, whereas a laissez-faire policy leading to segregation shaped by individual incentives is maximizing social welfare in the second-best case. Both these conclusions are valid in light of 'reasonable' concavity features of the individual utility function. An interpretation of our social welfare conclusions is that one should be cautious when arguing for an "always integration" policy. Our modelling approach provides an interesting complementary way of thinking about the arousal and maintenance of occupational segregation and wage and employment disparities to the existent economic frameworks. Eventually, an empirical test of the relevance of the mechanism we envisage here, relative to other models, would surely be needed; conditional on finding the suitable data, we postpone this exercise for future research.

Finally, Chapter 5 is the first study to investigate the impact of work environment health and safety practice on firm performance, and to reveal which firm-characteristic factors are associated with good workplace conditions. We merge three distinct Danish datasets: the longitudinal register matched employer-employee data, the longitudinal compulsory survey of firm business accounts and, the pivot dataset, a detailed cross-sectional representative survey of establishments on workplace safety and health conditions. This enables us to address standard econometric pitfalls such as omitted variables bias or endogeneity in estimating both production functions augmented with work environment indicators and aggregate employee characteristics, and firm mean wage regressions on the same firm aggregate employee characteristics. Our findings from the descriptive part of the analysis suggest, on the one hand, that few but robust factors are associated with a good work environment practice; these factors are the proportion of managers and respectively, courses with work environment content offered to all employees. This might suggest the importance of the high management involvement and respectively the role of the employee awareness in enhancing work conditions. On the other hand, from our main empirical analysis, the only work environment dimension compensated for by higher mean wages is found to be the "internal climate" at the firm, one of the "hardcore" physical dimensions. Finally, and crucially, the work conditions that contribute to higher firm productivity are having solved problems related to the same physical dimension, "internal climate" and respectively, "repetitive and strenuous work", while none of the other health and safety conditions are found to matter. This is a very interesting result in light of the strong promotion of firms' enhancement of psycho-social conditions at work by NGOs and policy decision makers for instance; however, before speculating on any policy advice, future research should compare our findings to results using data from other countries, ideally longitudinal in nature also in the workplace conditions dimension. This study retains the caveat of not exhaustively accounting for endogeneity of the workplace environment conditions in the production function estimation, problem present in all of the earlier (published) analogous methodological studies. Namely, in theory, some unobserved time-varying managerial ability or other unobserved time-varying firm-level heterogeneity, correlated with profits and not captured by lagged levels or changes in the proportion of managers, would still be able to bias our estimates. We believe however that in practice this is hardly likely to still be relevant.

Bibliography

Abowd, J.M. and D. Card (1989), "On the covariance of earnings and hours changes", *Econometrica* 57(2), 411-45

Abowd, J.M., F. Kramarz and D. Margolis (1999), "High wage workers and high wage firms", *Econometrica* 67(2), 251-333

Abraham, K.G. and H.S. Farber (1987), "Job duration, seniority, and earnings", *American Economic Review* 7(3), 278-97

Albaek, K., M. van Audenrode and M. Browning (1999), "Employment protection and the consequence for displaced workers: a comparison of Belgium and Denmark", in P.J. Kuhn (ed), *Loosing Work Moving On: Worker Displacement in International Perspective*, W.E. Upjohn Institute of Employment Research, Kalmazoo, Michigan

Albelda, R.P. (1986), "Occupational segregation by race and gender: 1958–1981," *Industrial and Labour Relations Review* 39, 404-11

Alesina, A. and E. La Ferrara (2000), "Participation in heterogeneous communities," *Quarterly Journal of Economics* 115, 847-904

Alesina, A. and E. La Ferrara (2002), "Who trusts others?," *Journal of Public Economics* 85, 207-34

Altonji, J.G. and R.A. Shakotko (1987), "Do wages rise with seniority?", *Review of Economic Studies*, 54, 437–59

Altonji, J.G. and N. Williams (1997), "Do wages rise with job seniority? A reassessment.", NBER Working Paper No. 6010

Altonji, J.G. and N. Williams (1999), "The effects of labor market experience, job seniority and job mobility on wage growth". in *Research in Labor Economics*, vol. 17, edited by S. Polacheck, JAI Press, Elsevier

Altonji, J.G. and R.M. Blank (1999), "Race and gender in the labor market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Volume 3, Amsterdam: Elsevier, 3143-3259

Altonji, J. G. and N. Williams (2005), "Do wages rise with job seniority? A reassessment", *Industrial and Labor Relations Review*, 58 (3), 370-97

Arrelano,. M. and S. Bond (1991), "Some tests of specifications for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, 58, 277-97

Arrelano, M. and O. Bover (1995), "Another look at the instrumental variable estimation of error-component models", *Journal of Econometrics*, 68, 29-51

Arrow, K.J. (1998), "What has economics to say about racial discrimination?," *Journal of Economic Perspectives* 12, 91-100

Arrow, K.J. and R. Borzekowski (2004), "Limited network connections and the distribution of Wages," FEDS Working Paper No. 2004-41, Federal Reserve Board, Washington

Askenazy, P. (2001), "Innovative workplace practices and occupational injuries and illnesses in the United States", *Economic and Industrial Democracy* 22(4), 485-516

Askenazy, P. and E. Caroli (2006), "Innovative workplace practice, innovative technologies and working conditions: Evidence from France", IZA Discussion Paper No. 2321

Bartel, A., R. Freeman, C. Ichniowski and M.M. Kleiner (2003), "Can a work organization have an attitude problem? The impact of workplaces on employee attitudes and economic outcomes", NBER Working paper 9987

Batt, R. (1999), "Work organisation, technology and performance in customer service and sales", *Industrial and Labor Relations Review*, 52(4), pp.539-61

Batt, R. (2002), "Managing customer services, human resources practices, quit rates and sales growth ", *Academy of Management Journal*, 45(3), 587-97

Beatson, M. and M. Coleman (1997), "International comparisons of the economic costs of work accidents and work-related ill-health", in J. Mossink and F. Licher (eds.), *Costs and Benefits of Occupational Safety and Health: Proceedings of the European Conference on Costs and Benefits of Occupational Safety and Health*, 1997, The Hague

Beaudry, P. and J. DiNardo (1991), "The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data", *Journal of Political Economy* 99(4), 665-88

Beller, A.H. (1982), "Occupational segregation by sex: Determinants and changes," *Journal of Human Resources* 17, 371-92

Benabou, R. (1993), "Workings of a city: Location, education, and production," *Quarterly Journal of Economics*, 108, 619-52

Bentolila, S. and G. Bertola (1990), "Firing costs and labour demands: how bad is eurosclerosis?", *Review of Economic Studies*, 57 (3), 381-402

Bertrand, M. and S. Mullainathan (2003), "Enjoying the quiet life? Corporate governance and managerial preferences", *Journal of Political Economy*, III(5), 1043-75

Bewley, T.F. (1999), *Why wages don't fall during a recession*, Cambridge MA: Harvard University Press

Bingley, P. and N. Westergård-Nielsen (2003), "Returns to tenure, firm-specific human capital and worker heterogeneity", *International Journal of Manpower*, 24(7), 744-88

Black, S.E. and L.M. Lynch (2001), "How to compete: The impact of workplace practices and information technology on productivity", *Review of Economics and Statistics*, 83(3), 434-45

Black, S.E. and L. M. Lynch (2004), "What's driving the new economy?: The benefits of workplace innovation", *Economic Journal*, 114, F97-F116

Bloom, N and J. van Reenen (2006), "Management practices, work-life balance, and productivity: A review of some recent evidence", *Oxford Review of Economic Policy*, 22(4), 457-82

Blundell, R. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87, 115-43

Blundell, R. and S. Bond (2000), "GMM estimation with persistent panel data: An application to production functions", *Econometric Reviews*, 19, 321-40

Bond, S., C. Nauges and F. Windmeijer (2005), "Unit roots: Identification and testing in micro panels", Cemmap working paper CWP07/05

Bound, J., and A. Krueger (1991),"The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?", *Journal of Labor Economics*, 9, 1-24

Bramoullé, Y. and G. Saint-Paul (2006), "Social networks and labour market transitions," mimeo, Université Laval and Université des Sciences Sociales de Toulouse

Brass, D.J. (1985), "Men's and women's networks: A study of interaction patterns and influence in an organization," *Academy of Management Journal* 28, 327-43

Breitung, J. and W. Meyer (1994), "Testing for unit roots in panel data: Are wages on different bargaining levels cointegrated?", *Applied Economics*, 26, 353-61

Brenner, M., D. Fairris, and J. Ruser (2004), "Flexible work practices and occupational safety and health: Exploring the relationship between cumulative trauma disorders and workplace transformation", *Industrial Relations: A Journal of Economy and Society* 43 (1), 242–66

Breusch, T.S. and A.R. Pagan (1979), "A simple test for heteroskedasticity and random coefficient variation", *Econometrica* 47, 1287-94

Brown, C. and J. Medoff (1989), "The employer size wage effect", *Journal of Political Economy*, 97, 1027–59

Buchinsky, M., D. Fougere, F. Kramarz and R. Tchernis (2005), "Interfirm mobility, wages, and the returns to seniority and experience in the US", IZA Discussion Paper

No. 1521

Buhai, I.S. and C.N. Teulings (2006), "Tenure profiles and efficient separation in a stochastic productivity model", Tinbergen Institute Discussion Paper 05-099/3, revised October 2006

Buhai, I.S. and M.J. van der Leij (2006), "A social network analysis of occupational segregation", Tinbergen Institute Discussion Paper 06-016/1

Buhai, I.S., M. Portela, C.N. Teulings and A. van Vuuren (2008), "Returns to tenure or seniority?", IZA Discussion Paper 3302

Buhai, I.S., E. Cottini and N. Westergård-Nielsen (2008), "The impact of workplace conditions on firm performance", mimeo Aarhus School of Business, University of Aarhus

Cabral, L.M.B. and J. Mata (2003), "On the evolution of the firm size distribution: facts and theory", *American Economic Review* 93, 1075–90

Calvó-Armengol, A. (2004), "Job contact networks," *Journal of Economic Theory* 115, 191-206

Calvó-Armengol, A. and M.O. Jackson (2004), "The effects of social networks on employment and inequality," *American Economic Review*, 94, 426-54

Calvó-Armengol, A. and M.O. Jackson (2007), "Networks in labor markets: Wage and employment dynamics and inequality," *Journal of Economic Theory*, 132, 27-46

Calvó-Armengol, A. and Y. Zenou (2005), "Job matching, social network and word-of-mouth communication," *Journal of Urban Economics*, 57, 500-22

Cappelli, P. and D. Neumark (2001), "Do "high performance" work practices improve establishment level outcomes?", *Industrial and Labor Relations Review*,54(4), 737-75

Cardoso and Portugal (2005), "Contractual wages and the wage cushion under different bargaining settings", *Journal of Labor Economics*, 23(4), 875-902

Caroli, E. and J. van Reenen (2001), "Skills and organisational change: evidence from British and French establishments in the 1980s and 1990s", *Quarterly Journal of Economics*, 116(4), 1449-92

Charles, M. and D.B. Grusky (2004), *Occupational ghettos: The worldwide segregation of women and men*, Stanford CA: Stanford University Press

Cohen, A. and L. Einav (2007), "Estimating risk preferences from deductible choices," *American Economic Review*, 97, 745-88

Cox, D. R. and H.D. Miller (1965), *The theory of stochastic processes*, London: Methuen

Currarini, S., M.O. Jackson and P. Pin (2008), "An economic model of friendship: Homophily, minorities and segregation," mimeo, Università di Venezia and Stanford University

Datcher, L. (1983), "The impact of informal networks on quit behavior," *Review of Eco-*

nomics and Statistics, 65, 491-95

Datcher Loury, L. (2006), "Some contacts are more equal than others: Informal networks, job tenure, and wages," *Journal of Labour Economics*, 24, 299-318

Dates, J. (1980), "Race, racial attitudes and adolescent perceptions of black television characters," *Journal of Broadcasting*, 24, 549-60

Dearden, L., H. Reed and J. Van Reenen (2006), "The impact of training on productivity and wages; Evidence from British panel data", *Oxford Bulletin of Economics and Statistics*, 68 (4), 397-421

Devine, T.J. and N.M. Kiefer (1991), *Empirical labor economics: The search approach*, New York: Oxford University Press

Dixit, A.K. (1989), "Entry and exit decisions under uncertainty", *Journal of Political Economy*, 97, 620-38

Dixit, A.K. and R.S. Pindyck (1994), *Investment under uncertainty*, Princeton University Press, Princeton, USA

Doeringer P.B. and M.J. Piore (1971), *Internal labor markets and manpower analysis*, Lexington: Heath Lexington Books

Dorman, P. (2000), "The economics of safety, health, and well-being at work: An overview", Infocus Program on Safe Work, International Labour Organisation

Dustmann, C. and C. Meghir (2005), "Wages, experience and seniority", *Review of Economic Studies* 72, 77-108

European Commission (2001), "Employment and social policies: A framework for investing in quality", Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions, COM(2001) 313 final, 20.06.2001

Farber, H.S. (1994), "The analysis of interfirm mobility", *Journal of Labor Economics* 12(4), 554-93

Farber, H.S. (1999), "Mobility and strategy: The dynamics of job change in labor markets", in O. Ashenfelter and D. Card, eds., *Handbook of Labour Economics*, Vol 3-B, Amsterdam, North-Holland

Feller, W. (1968), *An introduction to probability theory and its applications*, vol 2, Wiley Series in Probability and Mathematical Statistics, J. Wiley and Sons, Inc.

Fernandez, R.M. and M.L. Sosa (2005), "Gendering the job: Networks and recruitment at a call centre," *American Journal of Sociology*, 111, 859-904

Fisman, R., S. Iyengar, E. Kamenica, I. Simonson (2006), "Gender differences in mate selection: Evidence from a speed dating experiment," *Quarterly Journal of Economics*, 121, 673-97

Fontaine, F. (2008), "Do workers really benefit from their social networks?," *Louvain Economic Review*, forthcoming

Gertner, R. (1993), "Game shows and economic behavior: Risk taking on "Card Sharks"," *Quarterly Journal of Economics*, 108, 507-21

Gittleman, M., M. Horrigan, and M. Joyce (1998), "Flexible workplace practices: Evidence from a nationally representative survey", *Industrial and Labor Relations Review*, vol.52(1), pp.99-115

Gittel J.H., A. von Nordenflycht and T.A. Kochan (2004), " Mutual gains or zero sum? Labor relations and firm performance in the airline industry", *Industrial and Labor Relations Review*, 57(2), pp.163-80

Granovetter, M.S. (1973), "The strength of weak ties," *American Journal of Sociology*, 78,1360–80

Granovetter, M.S. (1995), *Getting a job: A study of contacts and careers*, 2nd edition, Chicago: University of Chicago Press

Granovetter, M.S. (2002), "A theoretical agenda for economic sociology," in M.F. Guillén, R. Collins, P. England and M. Meyer, eds., *The New Economic Sociology Developments in an Emerging Field*, New York: Russell Sage Foundation 35-59

Griliches, Z. and J. Mairesse (1998), "Production functions: The search for identification", in "Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium" (ed. S. Strom), 169-203, Cambridge University Press

Guthrie, J.P. (2001), "High-involvement work practices, turnover, and productivity: Evidence from New Zealand", *Academy of Management Journal*, 44 (1), pp. 180-90

Hasle, P. and N. Moller (2001), "The action plan against repetitive work- an industrial relation strategy for improving the work environment", *Human Factors and Ergonomics in Manufacturing*, 11, 131-44

Holzer, H.J. (1987), "Informal job search and black youth unemployment," *American Economic Review*, 77, 446-52

Horowitz, J.L. and S. Lee (2002), "Semiparametric estimation of a panel data proportional hazards model with fixed effects", Chapter 3, PhD Dissertation of Simon Lee, University of Iowa, 2002

Hosios, A.J. (1990), "On the efficiency of matching and related models of search and unemployment", *Review of Economic Studies*, 57, 279-98

Hraba, J. and G. Grant (1970), "Black is beautiful: A reexamination of racial preference and identification," *Journal of Personality and Social Psychology*, 16, 398-402

Huselid, M (1995), "The impact of human resource management practices on turnover, productivity, and corporate financial performance", *Academy of Management Journal*, 38(3), pp. 635-72

Ibarra, H. (1992), "Homophily and differential returns: Sex differences in network structure and access in an advertising firm," *Administrative Science Quarterly*, 37, 422-47

Ichiniowski, C., K Shaw and G. Prennushi (1997), "The effects of human resources management practices on productivity: A study of steel finishing lines", *American Economic Review*, 87(3), pp.291-311

Ichiniowski, C. and K. Shaw (1998), "The effects of human resources management systems on economic performance: An international comparison of US and Japanese plants", *Management Science*, 45(5), pp.704-21

Ioannides, Y.M. and L. Datcher Loury (2004), "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature*, 42,1056–93

Ioannides, Y.M. and A.R. Soetevent (2006), "Wages and employment in a random social network with arbitrary degree distribution," *American Economic Review Papers and Proceedings*, 96, 270-74

Jacobson, L.S., R.J. LaLonde and D.G. Sullivan (1993), "Earnings losses of displaced workers", *American Economic Review* 84(4), 685-709

Janod, V. and A. Saint-Martin (2004), "Measuring the impact of work reorganization on firm performance: Evidence from French manufacturing, 1995-1999", *Labour Economics* 11, 785-98

Jovanovic, B. (1979a), “Job matching and the theory of turnover”, *Journal of Political Economy*, 87(5), 972–90

Jovanovic, B. (1979b), "Firm-specific capital and turnover", *Journal of Political Economy* 87(6), 1246-60

Jovanovic, B. (1982), “Selection and the evolution of industry”, *Econometrica*, 50(3), 649-70

Kahneman, D., and A. Tversky (1979), "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 263-92

Kato, T. and M. Morishima (2002), "The productivity effects of participatory employment practices: evidence from new Japanese panel data", *Industrial Relations*, 41, 487-520

Katz H. C., T.A. Kochan and K.R. Gobeille (1983), " Industrial relations performance, economic performance and QWL programs: An interplant analysis", *Industrial and Labor Relations Review*, 37(1), pp. 3-17

Kenworthy, L. (2001), "Wage-setting coordination scores", mimeo, Emory University, Atlanta

Kijima, M. (2003), *Stochastic processes with applications to finance*, Boca Raton, FL, US: Chapman and Hall

King, M.C. (1992), "The evolution of occupational segregation by race and gender, 1940-1988," *Monthly Labor Review*, 115, 30-36

Koenker, R. (1981), "A note on studentizing a test for heteroskedasticity", *Journal of Econometrics* 17, 107-12

Kuhn, P. (1988), "A nonuniform pricing model of union wages and employment", *Journal of Political Economy*, 96, 473–508

Kuhn, P. and J. Robert (1989), "Seniority and distribution in a two-worker trade union", *Quarterly Journal of Economics*, 485-505

Lancaster, T. (1972), "A stochastic model for the duration of a strike", *Journal of the Royal Statistical Society A*, 135(2), 257-71

Lavezzi, A.M. and N. Meccheri (2005), "Social networks in labour markets: The effects of symmetry, randomness and exclusion on output and inequality," mimeo, Università di Pisa

Lazarsfeld, P.F. and R.K. Merton (1954), "Friendship as social process: A substantive and methodological analysis," in M. Berger, T. Abel and C. Page, eds., *Freedom and Control in Modern Society*, Princeton NJ: Van Nostrand,18-66

Lee, J. (2006), "American Idol: Evidence of same-race preferences?," IZA Discussion paper No. 1974

Leigh, J. P, S. Markowitz, M. Fachs, C. Shin, and P. Landrigan (1996), "Costs of occupational injuries and illnesses", National Institute for Occupational Safety and Health Report U60/CCU902886

Lindbeck, A. and D. Snower (1990), "Do cooperation and harassment explain involuntary unemployment?", *American Economic Review*, June, 80(3), 167-88

MaCurdy, T.E. (1982), "The use of time-series processes to model the error structure of earnings in a longitudinal data analysis", *Journal of Econometrics* 18, 83-114

Marsden, P.V. (1987), "Core discussion networks of Americans," *American Sociological Review*, 52, 122-31

Marsden, P.V. (1988), "Homogeneity in confiding relations," *Social Networks*, 10 , 57-76

Marschak, J. and W. H. Andrews (1944), "Random simultaneous equations and the theory of production,", *Econometrica*, 1263-97

Mayer, A. and S.L. Puller (2008), "The old boy (and girl) network: Social network formation on university campuses," *Journal of Public Economics*, 92, 329-47

McLaughlin, K. J.(1991), "A theory of quits and layoffs with efficient turnover," *Journal of Political Economy*, 99(1), 1-29

McPherson, M., L. Smith-Lovin and J.M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, 27, 415-44

Metrick, A. (1995), "A natural experiment in "Jeopardy!"," *American Economic Review*, 85, 240-53

Mincer, J. and B. Jovanovic (1981), "Labor mobility and wages", in S. Rosen (ed), *Studies in Labor Markets*, ch. 1, 21-63

Montgomery, J.D. (1991), "Social networks and labour-market outcomes: Toward an economic analysis," *American Economic Review*, 81, 1408-18

Montgomery, J.D. (1992), "Social networks and persistent inequality in the labour market," mimeo, Northwestern University, 1992

Mortensen, D.T. (1988), "Wages, separations, and job tenure: On-the-job specific training or matching?", *Journal of Labor Economics* 6(4), 445-71

Mortensen, D.T. (2003), *Wage dispersion: why are similar workers paid differently?*, MIT Press, Cambridge

Neal, D. (1995), "Industry-specific human capital: Evidence from displaced workers", *Journal of Labor Economics*, 14(4), 653-77

Nickell, S, L. Nunziata and W. Ochel (2005), "Unemployment in the OECD since the 1960. What do we know?", *Economic Journal*, 115, 1-27

OECD (2004), "Employment protection regulation and labour market performance", in: *OECD, Employment Outlook*, Chapter 2, 61-125

Osterman, P. (1994), " How common is workplace transformation and who adopts it?", *Industrial and Labor Relations Review*,47,173-88

Osterman, P. (2000), " Work reorganization in an era of restructuring : trends in diffusion and effects of employee welfare", *Industrial and Labor Relations Review*, 53, 179-96

Padavic, B.F. and I. Reskin (2002), *Women and men at work*, 2nd edition, Thousand Oaks: Pine Forge Press

Pettit, B. and J. Hook (2005), "The structure of women's employment in comparative perspective", *Social Forces*, 84, 779-801

Rees, A. (1966), "Information networks in labor markets," *American Economic Review*, 56, 559-66

Schelling, T.C. (1971), "Dynamic models of segregation," *Journal of Mathematical Sociology*, 1, 143-86

Shrum, W., N.H. Cheek and S.M. Hunter (1988), "Friendship in school: Gender and racial homophily," *Sociology of Education*, 61, 227-39

Simon, C.J. and J.T. Warner (1992), "Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure," *Journal of Labor Economics*, 10, 306-30.

Stern, S. (1997), "Simulation-based estimation", *Journal of Economic Literature* 35(4), 2006-39

Sutton, J. (1997), "Gibrat's legacy", *Journal of Economic Literature*, 35(1), 40-59

Sørensen, J.B. (2004), "The organizational demography of racial employment segregation," *American Journal of Sociology*, 110, 626-71

Tampubolon, G. (2005), "Fluxes and constants in the dynamics of friendships," ESRC Research Methods Programme Working Paper No. 25, University of Manchester

Topa, G. (2001), "Social interactions, local spillovers and unemployment," *Review of Economic Studies*, 68, 261-95

Teulings, C.N. and J. Hartog (1998), *Corporatism or competition? Labour contracts, institutions, and wage structures in international comparison*, Cambridge University Press

Teulings, C.N. and M.A van der Ende (2000), "A structural model of tenure and specific investments", Tinbergen Institute Discussion Paper 2000-9/3

Topel, R.H. (1991), "Specific capital, mobility and wages: wages rise with job seniority", *Journal of Political Economy*, 99(1), 145-76

Topel, R.H. and M.P. Ward (1992), "Job mobility and the careers of young men", *Quarterly Journal of Economics*, 107(2), 439-79

Van der Ende, M. (1997), *On the values of jobs and specific training*, PhD Dissertation at the University of Amsterdam and Tinbergen Institute

Weeden, K.A. and J.B. Sørensen (2004), "Two-dimensional models of sex segregation: Industries and occupations," in M. Charles and D.B. Grusky, eds., *Occupational ghettos: The worldwide segregation of women and men*, Stanford: Stanford University Press, 245-94

Whitmore, G.A. (1979), "An inverse Gaussian model for labour turnover", J*ournal of the Royal Statistical Society: A*, 142(4), 468-78

Wit, G. de (2005), "Firm size distributions. An overview of steady-state distributions resulting from firm dynamics models", *International Journal of Industrial Organization*, 23, 423-50

Wong, L.Y. (2003), "Why do only 5.5% of black men marry white women?," *International Economic Review*, 44, 803-26

Wooldridge, J. M. (2002), *Econometric analysis of cross section and panel data*, MIT Press

Zhang, P.G. (1998), *Exotic options* (2nd edition), World Scientific Publishing Co.

Zwick,. T. (2004), "Employee participation and productivity", *Labour Economics*, 11, 715-40

Samenvatting (Summary in Dutch)

Dit boek is samengesteld uit een viertal afzonderlijke studies die uiteenlopende onderzoeksgebieden binnen het brede vlak van de arbeidseconomie behandelen. Het doel van deze studies is om vooruitgang te boeken op de bestudeerde onderwerpen, waaronder zich enkele klassieke vraagstukken bevinden. Ik ben ervan overtuigd dat dit proefschrift deze doelstelling bereikt heeft: ik draag niet alleen bij aan innovatieve methoden, zowel op theoretisch als empirisch vlak, maar meer dan eens vind ik compleet nieuwe resultaten die vaak tegen de huidige economische gedachtengang ingaan. Daarnaast bevatten sommige bevindingen belangrijke beleidsaanbevelingen voor zowel de ondernemer als de beleidsmaker, waarbij wel moet worden opgemerkt dat het hier vooral de bedoeling is geweest om bij te dragen aan de economische literatuur in de hoop dat dit aanleiding geeft tot verder onderzoek. Te vaak hebben economen in het verleden getracht om nieuwe resulaten direct buiten de universiteit toe te passen terwijl het achteraf beter was geweest hier iets langer mee te wachten.

Hoofdstuk 2 ontwikkelt een nieuwe methode om loonprofielen te analyseren door gelijktijdig de ontwikkeling van lonen en de verdeling van senioriteit te modelleren. We ontwikkelen een theoretisch model dat kijkt naar de initiële baanspecifieke investeringen en de efficiënte onderhandelingen over het surplus van de baan. Hierbij veronderstellen we dat zowel de huidige baan als die van potentieel nieuwe banen een zogenaamde "random walk" volgen, zoals empirisch ook is vastgesteld. Deze opzet staat ons toe om reële optietheorie toe te passen. Daarmee zijn we in staat om de efficiënte beëindigingsregel van een aanstelling te berekenen. Ons model geeft een goede voorspelling van de feitelijke verdeling van de senioriteit. Omdat we de lonen slechts kunnen observeren aan het begin en aan het einde van een aanstelling, baseren we de empirische analyse op de verwachte loonstijging binnen de huidige baan, gegeven de lonen van potentieel concurrende banen. We toetsen ons model door gebruik te maken van een deel van de "Panel Study of Income Dynamics" en het blijkt dat de data heel goed verklaard wordt. Hierbij merken wij wel op dat lonen naar beneden toe star zijn, en daarom zou een verdere analyse hier ook rekening mee moeten houden. Dit is een eventuele uitbreiding voor toekomstige studies.

Een belangrijke conclusie uit ons werk is dat een model dat geen rekening houdt met de selectie-effecten van potentieel andere banen, een te hoge schatting van de loonprofielen oplevert. Deze overschatting is te vergelijken met die van OLS schatters in cross-sectie onderzoeken. We laten echter zien dat bijna de gehele opbrengst van senioriteit te danken is aan een dalende productiviteit in de overige banen in plaats van een stijgende productiviteit in de eigen baan. Indien dit gedeelte niet wordt meegenomen in de loonprofielen dan behoren onze schattingen tot de laagste schattingen die ooit gevonden zijn.

In hoofdstuk 3 bekijken we een tweetal empirische feiten waarbij we gebruik maken van een panel van werknemers en werkgevers in Denemarken en Portugal. Ten eerste worden de werknemers die als laatste aangenomen ook weer het eerst ontslagen worden, dat wil zeggen dat de "Last-In-First-Out" regel (LIFO) wordt toegepast. Ten tweede krijgen werknemers met een hogere senioriteit meer betaald. Hierbij vinden we dat het traditionele rendement dat mensen behalen uit de tijd die zij in de huidige baan zitten slechts als meetfout gezien kan worden die voortkomt uit de misspecificatie van de oorspronkelijke modellen. We verklaren deze verschijnselen door een dynamisch model van de onderneming te ontwikkelen waarbij we rekening houden met een stochastische goederenvraag en de aanwezigheid van opstartkosten bij het aannemen van een baan. Hierbij zijn laatstgenoemde gedefinieerd als eenmalige, onomkeerbare en baanspecifieke investeringen. Daarnaast staan we loononderhandelingen tussen werknemer en werkgever toe. Aanstellingen die beëindigd worden, dienen te voldoen aan de LIFO regel en onderhandelingen zijn efficiënt, wat inhoudt dat er geen surplus is op het moment dat een aanstelling beëindigd wordt. De LIFO regel leidt ertoe dat werknemers die al langer in de baan zitten een betere onderhandelingspositie verkrijgen, waardoor we een rendement zien op senioriteit. Een efficiënt personeelsbeleid houdt in dat de onderhandelingsmacht van de werknemers gelijk moet zijn aan hun aandeel in de baanspecifieke investeringen. De LIFO regel kan daarom gezien worden als een mogelijkheid om het eigendomsrecht van de baanspecifieke investeringen te beschermen. We beschrijven ook de invloed van ontslagkosten, ontslagbescherming en risicoaversie hierop. Een interessante gevolgtrekking van ons model is dat het rendement op senioriteit inhoudt dat de werknemer tot op zekere hoogte een aandeelhouder in haar eigen bedrijf is. Hoewel we de aanwezigheid van een rendement op senioriteit hebben aangetoond voor Denemarken en Portugal, kunnen we helaas niets zeggen over andere landen, en dan met name de Verenigde Staten. Er zijn echter sterke vermoedens dat dit het geval is. Deze suggestie wordt gewekt door het hoge rendement op senioriteit in de Verenigde Staten indien we dit vergelijken met Denemarken en Portugal. Daarnaast kan gesteld worden dat de LIFO regel in alle landen geldt en wat dat betreft kan specifieke regelgeving in Denemarken of Portugal meer gezien worden als een juridische vastlegging van ongeschreven regels en impliciete contracten die anders ook hadden gegolden.

In hoofdstuk 4 ontwikkelen we een simpel model waarin sociale netwerken de segregatie op de arbeidsmarkt kunnen verklaren, zij het op basis van ras, geslacht of etnische achtergrond. Nadat individuen een beslissing over hun carrière hebben genomen, verkrijgen zij een baan door middel van een netwerk van contacten die stochastisch maar wel met inteelt binnen de eigen groep worden gevormd. We tonen aan dat men zelfs met een kleine mate van inteelt er een stabiel evenwicht met segregatie van beroepen bestaat. Bovendien geldt dat, als het loonverschil tussen de beide beroepen niet zo groot is, volledige segregatie altijd een uitkomst is. Als het loonverschil groot is, dan kan volledige segregatie niet worden volgehouden, maar dan is een partieel segregatie-evenwicht houdbaar waarbij één groep zich volledig specialiseert in één beroep terwijl de andere groep zich verspreidt over de verscheidene beroepen. Bovendien is ons model in staat om langdurige werkloosheids- en inkomensverschillen te verklaren. Ook kijken we naar de implicaties van het model vanuit een planmatig oogpunt. Verrassend genoeg vinden we dat een sociaal optimaal beleid segregatie met zich mee brengt, zowel in het “first-best” als het “second-best” geval. In het ‘first-best’ optimum vinden we dat segregatie de voorkeur van de maatschappij als geheel heeft, terwijl een “laissez-faire” beleid het sociale welzijn in het “second-best” geval maximaliseert. Een mogelijke interpretatie van onze conclusies is dat men voorzichtig moet zijn met het beargumenteren van een “integratie is altijd beter” beleid. Onze benadering levert een interessante alternatieve denkwijze op wat betreft de blijvende verschijning van segregatie van beroepen en loon- en arbeidsongelijkheid. Uiteindelijk is er een empirische toets nodig op de relevantie van ons mechanisme ten opzichte van bestaande economische raamwerken. Onder de voorwaarde dat we geschikte data vinden, zouden we dit mogelijk in de toekomst kunnen uitvoeren.

Tot slot is hoofdstuk 5 de eerste studie die kijkt naar de invloed van werkomstandigheden (zoals veiligheid en gezondheid) op de resultaten van het bedrijf, waarbij we proberen te achterhalen welke factoren bepalend zijn voor een gezonde werkvloer. We voegen een drietal deense gegevensbestanden samen: een panel van werkgevers en werknemers, een panel van de balans en verlies- en winstrekening van bedrijven en het meest belangrijke gegevensbestand, een gedetailleerde cross-sectie op basis van een enquête over de veiligheid en de gezondheidsrisico’s op de werkvloer die onder een representatieve steekproef van bedrijfsvestigingen is genomen. Dit stelt ons in staat om standaard econometrische valkkuilen te omzeilen, zoals ontbrekende variabelen of endogeniteitsproblemen bij het gelijktijdig schatten van produktiefuncties en het uitvoeren van regressies op het gemiddelde loon binnen een bedrijf. Onze databeschrijvingen suggereren aan de ene kant dat een klein aantal robuuste factoren te maken heeft met een goede praktijk wat werkomgeving betreft: deze factoren zijn het aandeel aan managers, en het aanbod van cursussen voor werknemers waarbij aandacht aan de werkomgeving wordt geschonken. Mogelijk geeft dit het belang aan van zowel de betrokkenheid van het hoger personeel als van het werknemersbewustzijn bij het verbeteren van de werkomstandigheden. Aan de andere kant laat onze kernanalyse zien dat de enige werkomstandigheid die door een hoger loon wordt gecompenseerd, het "interne klimaat" in het bedrijf is. Tot slot, een belangrijk resultaat is dat de volgende werkomstandigheden de produktiviteit van een bedrijf bevorderen: het oplossen van problemen met betrekking tot het "interne klimaat" en "herhaaldelijk en inspannend werk", terwijl geen van de andere veiligheids- en gezondheidsrisico's van belang bleken te zijn. Dit is een erg interessant resultaat uit het oogpunt van de verbetering van de psychosociale werkomstandigheden, zoals aangemoedigd door bijv. NGO's en beleidsmakers. Voordat we gaan speculeren over beleidsadviezen, zou in de toekomst een vergelijkend landenonderzoek plaats moeten vinden, waarin bij voorkeur ook longitudinale data van de werkomstandigheden wordt beschouwd. De huidige studie heeft nog steeds last van een endogeniteitsprobleem dat net als in voorgaande studies niet volledig kan worden ondervangen. Er bestaat namelijk de mogelijkheid dat niet-waargenomen tijdsafhankelijke leidinggevende kwaliteiten of andere bedrijfsgerelateerde heterogene variabelen gecorreleerd zijn met winsten en niet opgepakt worden door het niveau of de verandering in het aandeel aan managers. Dit zou nog steeds voor onzuiverheden in de schattingen kunnen zorgen. We denken echter dat dit in de praktijk nauwelijks het geval zal zijn.

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:

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.
405. M.Â. DOS REIS PORTELA, Four essays on education, growth and labour economics.
406. S.S. FICCO, Essays on imperfect information-processing in economics.
407. P.J.P.M. VERSIJP, Advances in the use of stochastic dominance in asset pricing.
408. M.R. WILDENBEEST, Consumer search and oligopolistic pricing: A theoretical

and empirical inquiry.

409. E. GUSTAFSSON-WRIGHT, Baring the threads: Social capital, vulnerability and the well-being of children in Guatemala.
410. S. YERGOU-WORKU, Marriage markets and fertility in South Africa with comparisons to Britain and Sweden.
411. J.F. SLIJKERMAN, Financial stability in the EU.
412. W.A. VAN DEN BERG, Private equity acquisitions.
413. Y. CHENG, Selected topics on nonparametric conditional quantiles and risk theory.
414. M. DE POOTER, Modeling and forecasting stock return volatility and the term structure of interest rates.
415. F. RAVAZZOLO, Forecasting financial time series using model averaging.
416. M.J.E. KABKI, Transnationalism, local development and social security: the functioning of support networks in rural Ghana.
417. M. POPLAWSKI RIBEIRO, Fiscal policy under rules and restrictions.
418. S.W. BISSESSUR, Earnings, quality and earnings management: the role of accounting accruals.
419. L. RATNOVSKI, A Random Walk Down the Lombard Street: Essays on Banking.
420. R.P. NICOLAI, Maintenance models for systems subject to measurable deterioration.
421. R.K. ANDADARI, Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).
422. V. KARTSEVA, Designing Controls for Network Organizations: A Value- Based Approach
423. J. ARTS, Essays on New Product Adoption and Diffusion
424. A. BABUS, Essays on Networks: Theory and Applications
425. M. VAN DER VOORT, Modelling Credit Derivatives
426. G. GARITA, Financial Market Liberalization and Economic Growth
427. E. BEKKERS, Essays on Firm Heterogeneity and Quality in International Trade
428. H. LEAHU, Measure-Valued Differentiation for Finite Products of Measures : Theory and Applications
429. G. BALTUSSEN, New Insights into Behavioral Finance
430. W. VERMEULEN, Essays on Housing Supply, Land Use Regulation and Regional Labour Markets