



Department of Economics

Three Essays on Frictional Labor Markets

Georg Duernecker

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Florence
May 2010

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Three Essays on Frictional Labor Markets

Georg Duernecker

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Jury Members:

Prof. Lars Ljungqvist, Stockholm School of Economics
Prof. Salvador Ortigueira, EUI
Prof. Morten Ravn, University College London, Supervisor
Prof. Eran Yashiv, Tel Aviv University

© 2010, Georg Duernecker
No part of this thesis may be copied, reproduced or
transmitted without prior permission of the author

This dissertation is dedicated to

Rodion R. Raskolnikov, Salvo Montalbano and Harry Haller

and

my Mother

Acknowledgments

What is the difference between a *Bistecca alla Fiorentina* and a *Ph.D. Thesis*? ... There isn't any! The process of preparation of either requires considerable skill and near-artistic ability, plus a lot of painstaking handwork and scrupulous attention to detail. Of capital importance are undoubtedly a good measure of fine ingredients, as well as the patient hands and the culinary art of the chefs in charge of the preparation of either.

This thesis has greatly benefited from the culinary art of many chefs. One of which is my supervisor Morten O. Ravn to whom I am deeply grateful for his guidance, encouragement and insightful discussions over the past four years and also for his unconditional support on the academic job market. I would also like to thank my advisor Salvador Ortigueira for his straight-to-the-point, and thus very helpful, feedback and Ramon Marimon for taking the time to comment on my work. I also thank Per Krusell for his encouragement and his comments on my early dissertation work and for hosting me at Princeton University's Economics Department during the Fall of 2007, where I greatly benefited from the interaction with faculty members and Ph.D. students.

The administrative staff at EUI should not be left without mention. In particular, I thank Jessica Spataro, Marcia Gastaldo and Lucia Vigna for their work and logistic support and Thomas Bourke for his great enthusiasm and dedication to his work. I also gratefully acknowledge the financial support from the Austrian Federal Ministry of Science and Research.

Over the past four years I have had the pleasure and the privilege to meet a number of very fine people with whom I have spent many cheerful moments over countless caffè, gelati and pizze and who also accompanied me through the more difficult stages of my Ph.D. experience. Among those are Marco Lenzi, Afrodite Kero, Marta Arespa, Cosimo Pancaro, Alicia Perez-Alonso, Konrad Smolinski, Isabella Sileoni, Nicolas Serrano-Velarde and Maren Frömel.

Kathia Serrano has been by my side for much longer than the official calendar says. She is the most precious thing I am taking with me from Florence.

Contents

I	Introduction	1
II	Chapters	5
1	Technology Adoption, Turbulence and the Dynamics of Unemployment	6
1.1	Introduction	7
1.2	The Facts	11
1.2.1	Unemployment	11
1.2.2	The Arrival Rate of New Technologies	13
1.2.3	Technology and Growth Gap	14
1.3	The Model	17
1.3.1	Vintage Technology and Skills	18
1.3.2	Unemployment	20
1.3.3	The Labor Market	21
1.3.4	Government	22
1.3.5	Value functions and the firm's updating problem	22
1.4	Solving the model	26
1.5	Parametrization and Calibration	27
1.6	Results	30
1.6.1	The pre-1975 period	30
1.6.2	The post-1975 period	33
1.7	Sensitivity Analysis	39
1.7.1	The role of unemployment insurance	39
1.7.2	Do specific skills matter? The effects on wages and unemployment . .	41
1.8	Conclusion	43
	Bibliography	45

2	Informational Frictions and the Life-Cycle Dynamics of Job Mobility	50
2.1	Introduction	51
2.2	Empirical Facts	59
2.2.1	Individual job stability and turnover	59
2.2.2	Unemployment	64
2.3	The Model	65
2.3.1	Workers	66
2.3.2	Firms	67
2.3.3	Match Formation	68
2.3.4	Value Functions	69
2.3.5	Equilibrium	72
2.4	<i>A Simplified Version</i>	75
2.4.1	Equilibrium	76
2.4.2	Threshold signals and threshold beliefs	77
2.4.3	Aggregate Outcomes	78
2.5	Solving the Full Model	81
2.6	The Estimation Strategy	82
2.7	The Data	87
2.8	Estimation Results	92
2.8.1	Coefficient estimates of the auxiliary model	92
2.8.2	Estimating the parameters of the structural model	98
2.8.3	Discussion of the Results	104
2.9	Conclusion	114
	Bibliography	115
2.10	Appendix A	118
2.10.1	Proof of Proposition 2.1	118
2.10.2	Proof of Proposition 2.4	119
2.11	Appendix B: Threshold signals and threshold beliefs	120
2.12	Appendix C: Computing aggregate variables in the simple model	121
2.12.1	The type-specific unemployment rate	121
2.12.2	Match destruction	122
2.13	Appendix D	123
2.13.1	The joint surplus function	123
2.13.2	Wage formation	123
2.14	Appendix E	124

2.15	Appendix F: Accuracy of the estimation results	125
3	On the Determinants, Nature and Consequences of Job Separation	126
3.1	Introduction	127
3.2	The Empirical Framework	129
3.2.1	Addressing Question 1: The determinants of job separation	129
3.2.2	Addressing Question 2: Investigating the nature of job separations . .	131
3.3	The Data	132
3.4	Estimation Results	136
3.4.1	Results for Question 1: The determinants of job separation	136
3.4.2	Results for Question 2: Investigating the nature of job separations . .	143
3.4.3	Results for Question 3: Reemployment Wages	146
3.5	Discussion and Conclusion	150
	Bibliography	152

Part I

Introduction

Introduction

The labor market is central to many issues in economics, including business cycles, unemployment, inequality, education, and growth. Moreover, it is the largest single market in most economies and it is fundamental in determining individual and household well-being. Therefore, a good understanding of the many phenomena that we observe in modern societies requires knowledge of the functioning of the labor market. It is the aim of this dissertation to improve and to deepen this knowledge and it does so by investigating three important labor market phenomena. In what follows I will provide a brief description of each of the chapters.

In the **first chapter**, I address the divergence of unemployment rates between the U.S. and Europe. The motivation for the analysis is based on the following observation: After low levels of unemployment until the late 1970s, European unemployment became high relative to that in the United States. Labor markets in Europe started to deteriorate at a time when there was a substantial acceleration in the arrival of new technologies as measured by the rate of capital-embodied technical change. There is convincing empirical evidence, some of which is provided by Oliner and Sichel (2000), Jorgenson and Stiroh (2000) and van Ark et al. (2002), indicating that certain economies in Europe have been lagging behind the U.S. (and other European economies) in the adoption and usage of new technologies.

In the first chapter I argue that the coexistence of a technology deficit - resulting from slack technology adoption - and the divergence of unemployment rates across economies are not coincidental. I develop a theoretical model and show that the speed with which firms in an economy adopt new technologies is a key determinant for how the economy's labor market reacts to an acceleration in capital-embodied growth. I find that the observed cross-country differences in firms' technology adoption can account for a large part of the observed divergence of unemployment rates across economies. The mechanism I propose can explain both (a) the divergence of unemployment rates between the major European countries and the U.S. and (b) the observed variation of unemployment rates across European economies. Previous work in this field could explain just the first feature but failed to account also for the divergence of unemployment rates across European economies.

In the **second chapter**, I investigate the life-cycle dynamics of individual job mobility. The central question I pose is: Why does individual job mobility systematically differ across age groups? What we observe in the data is that young individuals typically change jobs very frequently and retain each new job just for a short period of time. By contrast, experienced

workers tend to hold stable jobs, but more importantly, they are also less likely to separate from a new job than the young. This pattern is a common characteristic of all labor markets in OECD countries, though it is particularly pronounced in the U.S.

The literature considers the initially high, but declining, turnover for the young to be the outcome of a search process in which workers experiment with jobs in order to find the right match. This explanation is incomplete as it implies that each prime-age worker, who is laid off from a long-term job and seeks to find a new occupation would, necessarily, have to go through the same wasteful search process as at the beginning of the career, which is empirically not the case.

The objective of this chapter is to contribute to the understanding of the observed pattern of individual job mobility along the life-cycle. To this end, I first present new empirical evidence documenting the life-cycle dynamics of job mobility in the U.S. labor market. Second, in order to account for the observed pattern, I construct a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. I model the imperfection as a noisy signal about the match quality that a firm and worker observe when they first meet. The key ingredient is that the imperfection is assumed to be worker-specific and in particular, it is linked to an individual's previous labor market history. Thus, the informativeness of the signal differs across workers as each of them has built up an idiosyncratic labor market history. The structural parameters of the model are estimated using data from the National Longitudinal Survey of Youth 1979 (*NLSY 79*) covering the period between 1979 and 2006.

The estimated model can account very well, qualitatively and also quantitatively, for the observed pattern of job mobility along the life cycle. The model predicts that (a) young individuals are substantially more likely to separate from a new job than experienced workers, (b) the hazard rate of separating from a new job is increasing in the length of the retention period and (c) the hazard rate of separating from any given job is declining with tenure, for all age groups. Moreover, the model matches the observed extent of job turnover early in the career as well as the gradual decline in later stages. The implied path for the cumulated number of jobs, at each age, as a fraction of the career total is consistent with what we observe in the data.

The **third chapter** empirically sheds some light on the causes and consequences of job separation. In particular, it addresses the following three questions : (a) What are the factors determining the separation hazard of employment relationships? (b) Which employer-employee matches are more likely to dissolve due to a layoff rather than a voluntary quit?

(c) What are the effects of a voluntary quit and a layoff, respectively, on re-employment wages? Those questions are addressed using data on U.S. white males from the National Longitudinal Survey of Youth 1979. The amount of labor reallocation taking place in modern economies is typically very large. According to Fallick and Fleischman (2004), 6.7% of all U.S. employee-employer matches are dissolved in an average month. At the same time, 4.3% of the civilian noninstitutional population aged 16 and over began to work in a new job. Many of the newly created matches survive only for a short period of time. A significant proportion, however, is more durable and turns into medium and long-term jobs. This raises an obvious question, namely: what are the underlying factors and determinants that make some of the new matches likely to survive longer than others. Job survival is not a purely random process and we may presume that it is in fact governed by a given (possibly very large) set of determinants.

In this chapter, I identify worker- and match-specific factors that are causal for the observed systematic differences in the actual duration of employment spells. An employment spell can end due to a quit or a layoff. Both have very different implications for an individual's economic well-being. A layoff is typically perceived as an adverse (and often unexpected) economic shock whereas, a quit is considered as something positive, as it often leads to an immediate (and foreseen) transition to another job and thereby to an improvement in an individual's economic situation. I also provide a set of factors which help to assess whether a given match is more likely to dissolve due to a layoff rather than a voluntary quit? And lastly, I very briefly analyze the effects on re-employment wages of a voluntary quit and a layoff.

The chapter offers a set of answers to the aforementioned questions, some of which confirm previous findings in the literature, whereas others provide new facts and insights and also pose a challenge to conventional theoretical labor market models, such as the standard job ladder model. I intend to use the controversial elements as a point of departure for future research on individuals' labor market dynamics.

Part II

Chapters

Chapter 1

Technology Adoption, Turbulence and the Dynamics of Unemployment

Abstract

The divergence of unemployment rates between the U.S. and Europe coincided with a substantial acceleration in capital-embodied technical change in the late 70s. Furthermore, evidence suggests that European economies have been lagging behind the U.S. in the adoption and usage of new technologies. This paper argues that the pace of technology adoption plays a fundamental role in the manner in which a labor market reacts to an acceleration in capital-embodied growth. The framework proposed offers a novel explanation for the divergence of unemployment rates across economies that are hit by the very same shock (i.e. the acceleration in embodied technical change) but differ in their technology adoption. Moreover, the results of the paper challenge the popular - but controversial - view that high European unemployment is the result of institutional rigidities by claiming that institutions are not the principal cause, per se, but rather that they amplify certain forces that promote the emergence of high unemployment.

JEL Classification: J24, J64, O33

Keywords: Unemployment, Matching, Turbulence, Technology Choice, Capital-embodied Technical Change, Skill Loss

1.1 Introduction

After low levels of unemployment until the late 1970s, European unemployment became high relative to that in the United States. Labor markets in Europe started to deteriorate at a time when there was a substantial acceleration in the arrival of new technologies, as measured by capital-embodied technical change. Documented by Gordon's (1990) influential work on the quality-adjusted price of capital, and more recently by Cummins and Violante (2002), the rate of change in the relative price of new capital investments in the U.S. has decreased substantially from -2% before the mid-70s to -4.5% in the 1990s, suggesting an acceleration in capital-embodied technical change. There is convincing empirical evidence, some of which is provided by Oliner and Sichel (2000), Jorgenson and Stiroh (2000) and van Ark et al. (2002), indicating that certain economies in Europe have been lagging behind the U.S. (and other European economies) in the adoption and usage of new technologies. This is reflected by a persistent technology gap exhibited by those countries - as measured by labor productivity growth in the manufacturing sector, the share of information and communication technologies (ICT) in investment and its contribution to output growth.

This paper argues that the coexistence of a technology deficit - resulting from slack technology adoption - and the divergence of unemployment rates across economies is not coincidental. I show that the speed with which firms in an economy adopt new technologies is a key determinant for how the economy's labor market reacts to an acceleration in capital-embodied growth. I propose a framework in which cross-country differences in firms' technology adoption can account for a large part of the observed divergence of unemployment rates across economies. The mechanism in this paper can explain both (a) the divergence of unemployment rates between the major European countries and the U.S. and (b) the observed variation of unemployment rates across European economies. Previous work in this field could explain only the first feature but failed to account for the divergence of unemployment rates across European economies.

Moreover, the results of the paper challenge the popular - but controversial - view that high rates of European unemployment are the result of labor market institutions by claiming that institutions are not the principal cause of high unemployment, per se. Rather, they represent a tendency to amplify certain forces that promote the emergence of high levels of unemployment¹.

¹The controversy comes from the fact that the institutions that are held responsible were also present in the 1960s, yet in the 1960s unemployment was much higher in the U.S. than in Europe, see e.g. Blanchard and Wolfers (2000).

The analysis in this paper is based on a labor market matching model that is augmented by an endogenous technology choice by firms and a skill accumulation technology for workers. The channel through which firms' technology adoption affects an economy's labor market performance works as follows. The frequency with which firms adopt new technologies determines how long a firm's production technology is kept in operation. If updating is fast (slack) technologies are replaced rather (in)frequently and therefore workers stay with a certain technology for a relatively short (long) time. Constant advancements at the technology frontier gradually render each existing technology in the economy obsolete. Consequently, firms with infrequent technology updating operate technologies that are, on average, far-off the frontier. Workers that are attached to those obsolete technologies will possess a production knowledge - that is the set of skills and abilities needed to operate the specific technology - that is also relatively outdated with respect to the production knowledge associated with the frontier technology². If, on the other hand, firms update rather infrequently then workers' production knowledge is, on average, less obsolete as the firms' technologies stay closer to the frontier.

Therefore, the frequency with which firms adopt new technologies determines the average degree of workers' human capital obsolescence in an economy and thereby, affects the expected job creation costs for firms. This is due to the fact that a firm, when hiring a new worker, needs to provide costly training in order to overcome the worker's skill obsolescence - i.e. to make the worker's human capital compatible with the firm's technology³. These costs are increasing in the degree of obsolescence. Hence, more obsolete skills require more training which will drive up expected costs of creating a job.

An acceleration in the rate of embodied technical change - similar to that observed in the mid 1970s - causes the human capital of unemployed workers to depreciate more quickly, as it makes the gap between workers' current level of knowledge and that at the frontier widening more rapidly. This raises the expected training costs of firms when hiring workers. Higher training / job creation costs reduce the net present value of a job for a firm and discourages job creation. The negative effect on job creation is stronger - and the associated rise in unemployment is higher - in economies in which firms update their technologies less frequently.

²The process of technical change gradually renders a worker's human capital obsolete by driving a gap between a worker's current level of knowledge and that required at the technology frontier.

³The vintage specificity of workers' production knowledge implies that if a worker gets matched with a technology different from her current one then there will be a skill mismatch - in the sense that the worker's current level of knowledge is not fully compatible with the respective technology.

This is due to the fact that in those economies workers stay with the same technology for a longer period of time, increasing the average degree of human capital obsolescence among workers entering the unemployment pool.

In recent years, economists have offered numerous explanations for the emergence of high European unemployment in the late 1970s, involving factors such as overly generous welfare systems, slow TFP growth or capital market imperfections⁴. One particularly influential strand in the literature emphasizes the interaction of macroeconomic shocks and labor market institutions as the main driving force for high levels of European unemployment. Key references include Ljungqvist and Sargent (1998, 2007), Marimon and Zilibotti (1999) and Hornstein, Krusell and Violante (2007). The framework proposed by Ljungqvist and Sargent (1998) is the first rigorous attempt to study the shock-policy interaction within a calibrated model. A related explanation is offered by Marimon and Zilibotti (1999). The line of argument proposed by these authors is as follows. European unemployment increased due to reduced workers' incentives to exit unemployment. Workers in Europe prefer to collect generous unemployment benefits rather than to work for a low wage. Wages are low because the technology shock has made workers' skills obsolete - as in Ljungqvist and Sargent (1998) - or made it increasingly difficult to match with existing vacancies - as in Marimon and Zilibotti (1999). The mechanism in those papers operates primarily through the labor supply side. Ljungqvist and Sargent (2007) is a refinement that considers a matching framework in which firms adjust labor demand in the aftermath of the shock. The shock considered in Ljungqvist and Sargent (1998, 2007) refers to a general change in the economic environment, i.e. an increased degree of economic turbulence, rather than an explicit shock to technological change.

Recently, a number of economists have pointed to the potential significance of embodied technical change in order to explain the differences in labor market outcomes across countries. Hornstein, Krusell and Violante (2007) was the first work to highlight the interaction between shocks to capital-embodied technical change and labor market institutions. In their model, an increase in embodied technical change - such as that observed in the mid-70s - leads to a sharp reduction in firms' labor demand in a (European-type) welfare state economy, whereas it has only mild effects on labor demand in a (U.S.-type) laissez-faire economy. Consequently, unemployment rises by much more in the welfare state.

⁴See Nickell (2003) for a recent survey of research on the issue of European unemployment. Blanchard (2005) is an excellent assessment of the state of the contemporary literature regarding the European unemployment question.

However, these papers suffer from the following drawback. The mechanisms proposed are designed to reproduce the movements of an average European unemployment rate but they fail to account for the large heterogeneity of unemployment rates across European countries. In fact, Blanchard (2005) suggests that discussing "European unemployment" is misleading since high average European unemployment reflects high unemployment in four large continental countries (Germany, Italy, Spain and France) whereas unemployment is low (and comparable to the U.S. rate) in many other European countries. Arguably, a theory that addresses the European unemployment experience but fails to explain the large heterogeneity of labor market outcomes across European economies clearly conflicts with an essential aspect of actuality and is likely to disregard relevant factors.

The underlying determinants for the observed divergence are unlikely to be of an institutional nature. Even though changes in labor market institutions can account for parts of the rise in European unemployment - as shown by Nickell et al. (2005) - a sizable fraction remains unexplained. This paper proposes a framework in which ex-ante differences in firms' technology adoption across countries can account for a large part of the observed divergence of unemployment rates across economies. It does not rely on differences in institutional settings but rather stresses the role of the particular technological environment in explaining the unemployment experiences of Europe and the U.S.

The model builds upon the matching framework by Mortensen and Pissarides (1998). It augments the standard model by an endogenous technology choice by firms and a skill accumulation technology for workers. However, the model of Mortensen and Pissarides (1998) - as Hornstein et al. (2007) - display one stark feature: workers are not constrained by any skill requirements when moving across technologies of different levels of advancement. In other words, individuals in those models can switch to more advanced technologies without any extra cost. It is, therefore, equally costly for firms to hire a worker coming from a high-tech or low-tech job⁵. However, one might conjecture that technologies differ with respect to the set of skills and abilities required to operate them. Newer and more advanced technologies require a different set of skills than older and less advanced ones. Arguably, in a framework that considers capital-embodied technical change, the issue of workers' skill obsolescence is of particular relevance since the set of skills needed (e.g. for the most advanced technology) is subject to change over time.

⁵The only cost of the labor reallocation process is due to search. The explicit acquisition of skills, however, is generally uncoupled from a worker's previous experience.

The framework presented in this paper explicitly models workers' skill dynamics. It can account for increased human capital obsolescence caused by a shock to embodied technical change. This will be an important determinant for the dynamics of job creation in the aftermath of a shock. Ljungqvist and Sargent (1998, 2007) also consider a form of human capital obsolescence but the mechanism proposed in this paper is significantly richer. Unlike Ljungqvist and Sargent (1998, 2007), which are not explicit about the underlying economic mechanism, the degree of skill obsolescence in this paper is endogenous as it is driven by firms' technology choices. Thereby, the paper adds microfoundation to the turbulence approach forwarded by Ljungqvist and Sargent (1998). In that way, the mechanism also provides a rationale for higher turbulence (skill obsolescence) since it directly relates increased human capital obsolescence to the observable acceleration in capital-embodied technical change in the mid 1970s.

The remainder of the paper is organized as follows. Section 1.2 presents a variety of facts that motivate the analysis. Section 1.3 presents the theoretical model. Sections 1.4 and 1.5 briefly discuss the parametrization and calibration of the model and sketch the algorithm used to solve the model. Section 1.6 presents and discusses the results and Section 1.7 performs a variety of sensitivity checks. Section 1.8 concludes.

1.2 The Facts

1.2.1 Unemployment

In the postwar period up until the late 1970s, unemployment in Europe was low relative to that in the U.S. The data summarized in Table 1.1 shows that during the whole period up until the 1980s, unemployment in the U.S. was significantly higher than in Europe. In the 1960s and early 70s, the average unemployment rate in European was around 2.5% whereas the U.S. figure was around 5%. However, the picture changed dramatically after the mid-1970s. Unemployment in Europe experienced a sharp and persistent increase up to a level of around 9% whereas U.S. unemployment rose by much less. This rise in Europe, however, was not homogenous across economies. In some countries the increase in unemployment was much less pronounced (as in Sweden, Denmark or Austria) than in others (such as France, Germany or Italy). Therefore, the use of the average rate of unemployment is actually very misleading. Table 1.1 reveals that there is in fact substantial variation of unemployment rates across European economies. Over the period 2000 – 06, for instance, six out of the twelve European labor markets in Table 1.1 produced unemployment rates that are just slightly

above or even below the U.S. rate. This implies that when we exclude some of the major European countries, in particular Germany, France and Italy, the *European unemployment puzzle* vanishes⁶. High unemployment is not, therefore, a phenomenon that is specific to Europe, per se, but rather to certain countries.

	1960 - 69	1970-74	1975-79	1980 - 89	1990 - 99	2000 - 06
U.S.	4,78	5,38	7,04	7,27	5,71	5,11
Euro Area	2,16	2,62	4,76	8,36	9,55	8,43
Austria	2,13	1,30	1,56	2,70	3,87	4,36
Belgium	1,93	2,04	5,96	9,54	8,47	7,74
Denmark	1,11	1,16	5,22	6,67	6,93	4,71
Finland	2,20	2,40	5,16	4,79	11,87	8,84
France	1,75	2,68	4,84	8,54	10,60	9,19
Germany	0,65	0,90	3,12	5,83	6,29	8,46
Italy	4,86	5,42	6,46	8,39	10,25	8,39
Netherlands	0,84	1,98	5,62	8,16	5,41	3,53
Norway	0,00	1,60	1,90	2,81	4,80	3,99
Spain	2,39	2,92	6,12	15,44	15,89	10,29
Sweden	1,70	2,24	1,86	2,59	7,21	5,96
UK	1,61	2,44	4,54	9,45	8,00	5,01

Source: European Commission, Annual Macroeconomic Database (AMECO)

Table 1.1: Average annual unemployment, in %, by subperiod

A distinguishing feature of the U.S. labor market is its fluid nature. The average duration of unemployment is low relative to many European countries, as evidenced by a rapid reallocation of labor across sectors⁷. Table 1.2 shows that in the period 2000 – 04, the fraction of unemployed being jobless for less than one month is 37.22% in the U.S. while it is around 5% in Germany, France, Spain and Italy. In contrast, only 9% of the unemployed in U.S. stay out of work for more than one year whereas the number for Germany, France, Spain and Italy ranges from 40% to 58%. Evidently, high unemployment rates in some European countries are the result of a massive share of long-term unemployed.

⁶The so-called "*European Unemployment Puzzle*" refers to high and persistent rates of unemployment in Europe relative to that in the United States.

⁷Farber (1999) finds that in the U.S. half of all new jobs end in the first year and at any point in time about 20% of workers have been with their current employer for less than one year.

	Belgium	Germany	Denmark	Spain	France	UK	Italy	Sweden	U.S.
< 1 month	7.4	5.76	22.18	5.34	4.48	16.2	5.1	21.34	37.22
[1, 3) months	11.18	11.7	18.96	16.78	18.46	23.98	8.92	22.56	30.8
[6, 12) months	17.12	16.64	18.28	18.84	19.48	15.68	15.66	16.16	8.06
≥ 12 months	50.7	50.32	20.88	41.86	39.7	24.66	58.36	21.28	9.02
Source: OECD, Numbers indicate the fraction of unemployed by duration, period: 2000-04									

Table 1.2: Average duration of unemployment

1.2.2 The Arrival Rate of New Technologies

There is evidence, some of which is provided by Cummins and Violante (2002), Greenwood and Yorukoglu (1997) and Pakko (2002), that the rate of arrival of new technologies has increased quite substantially since the late 1970s. Cummins and Violante (2002) constructs an aggregate index of investment-specific technological change. They find that average annual growth rates were stable around 4% until the late 1970s but then there was a sharp acceleration in the 1980s that lead to annual growth rates of more than 6% in the 1990s. As argued by Hornstein and Krusell (1996) and Yorukoglu (1998), an increase in the arrival rate of new technologies has important consequences for the process of technology adoption. A higher rate of technological change means that new technologies, which differ substantially in their characteristics from existing ones, are introduced at a faster rate. This raises the issue of compatibility problems between consecutive vintages. The improved technology embodied in new capital changes the technological standards and decreases the compatibility between old and new vintages. Yorukoglu (1998) argues that the more advanced the new technology is relative to the existing one, the lower is the initial experience with the new production technology.

This implies that, as the rate of technological change increases, agents will be less familiar with the new technology and it will be more costly to adopt. Therefore, in times of rapid technological change we should see an increase in the technology gap and a rise in the adoption costs per unit of investment. Regarding the former, Cummins and Violante (2002) finds that the technology gap in the U.S. (which they define as the gap between the productivity of the best technology and the productivity of the average practice in the economy) was 15% in 1975. In 2000, the figure had jumped to 40%. These findings suggest that firms were not able to keep track with the accelerated process of technical change, or in other words, the increased speed of technological change has outpaced firms' technology updating.

Concerning technology updating costs, Bessen (2002) provides evidence that adjustment costs rose sharply during 1974-83 and more than doubled from the early 1960s to the late

1980s⁸. He finds that the costs associated with adopting a new technology amounted to 0.35\$ per dollar of investment in the period 1961 – 73. In the period 1974 – 83, adjustment costs rose sharply to 0.79\$ per dollar of investment and peaked at 0.90\$ in 1984 – 88. As a result, the adoption costs as a percentage of aggregate output increased from 2.4% in 1973 to 6.5% in 1983. Bessen (2002) argues that the rise in costs is specifically associated with a switch in firms' investment towards new technologies.

1.2.3 Technology and Growth Gap

Economic growth in Europe was strong until the 1980s but became weaker in the subsequent decades. As a result, a persistent growth gap, in both GDP growth and labor productivity growth, between the U.S. and most European countries has emerged since the 1980s. By taking data on relative manufacturing output per person, Scarpetta et al. (2000) shows that the productivity level for Germany and other European countries was converging toward the U.S. level until the 1980s but has since diverged.

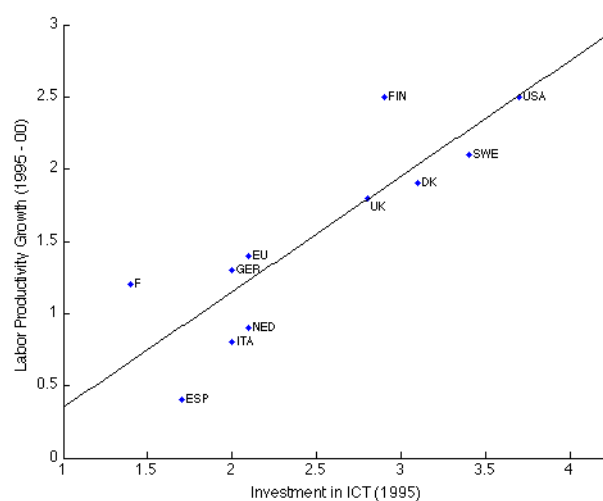


Figure 1.1: Investment in new technologies and LPD growth, Data is from van Ark et al. (2002) and Timmer et al. (2003)

At the same time, several Europe economies have lagged behind the U.S. in terms of adoption and usage of new technologies. Timmer et al. (2003) reports that many EU countries have been seriously lagging behind the U.S. in the share of ICT investment in GDP.

⁸He estimates technology adoption costs in the U.S. manufacturing sector for the period 1961-96.

Consequently, IT capital stocks are much lower in those countries. It is a well established fact that slower rates of ICT investment are key in explaining the poorer European productivity performance. Figure 1.1 illustrates the strong positive relation between the level of investment in new technologies and labor productivity growth.

Lower investment rates in ICT mean that newer technologies have been adopted less forcefully. In fact Oliner and Sichel (2000) and Jorgensen and Stiroh (2000) provide evidence that, to a large degree, the U.S.-EU productivity gap can be traced back to the delayed adoption of new technologies in Europe. This finding is confirmed by a number of studies, including Daveri (2002), Colecchia and Schreyer (2002) and van Ark et al. (2002). Further, Daveri (2002) and van Ark et al. (2002), among others, find that the diffusion of new technologies in Europe is following a similar pattern to that observed in the U.S., albeit at a considerably slower pace. This pattern is clearly observable in Table 1.3. It shows that ICT investment intensities were increasing in all countries over time but (a) most European countries started investing in ICT with a significant delay and (b) the gap between the U.S. and most European economies has not narrowed greatly.

	1980	1985	1990	1995	2001
U.S.	2.5	3.4	3.3	3.7	4.2
EU	1.2	1.9	2.2	2.1	2.6
Denmark	1.5	2.3	2.9	3.1	3.6
Finland	1.1	1.5	1.9	2.9	4.3
France	1.0	1.4	1.5	1.4	2.1
Germany	1.3	2.2	2.4	2.0	2.5
Italy	1.5	1.9	2.3	2.0	2.5
Netherlands	1.6	1.9	2.4	2.1	2.9
Spain	0.9	1.5	2.5	1.7	2.1
Sweden	1.6	2.5	2.7	3.4	4.7
United Kingdom	0.8	1.6	2.3	2.8	3.0
Source: Timmer et al. (2003)					

Table 1.3: ICT investment as a share of GDP

However, the lagging technology diffusion is not a feature that applies to Europe in general. As Figure 1.1 illustrates there is in fact a remarkable cross-country heterogeneity with respect to investment in new technologies. Several European countries - most notably Sweden, Denmark and the UK - feature *ICT* investment rates that are comparable to the U.S. rate whereas other countries - such as Spain, Italy or France - seem to be lagging behind.

From this we may conclude that European economies seem to differ substantially with respect to the speed with which firms invest in new technologies. Consequently, the technology gap will also be very different across countries.

A question that is of interest in this context is why technology adoption is so different across economies. There exists a bulk of empirical studies, for instance McGuckin and van Ark (2001) and McGuckin et al (2005), arguing that structural impediments in product markets hamper the successful implementation of new technologies across industries in Europe. These barriers mostly come in the form of burdensome regulations. Regression estimates by Nicoletti and Scarpetta (2003) suggest that strict product market regulations that curb competition hinder the adoption and diffusion of new technologies and thus have a negative effect on productivity. More evidence provided by Gust and Marquez (2002) suggests that countries with more burdensome regulatory environments tend to adopt new technologies more slowly and also have slower productivity growth. These studies argue that as adoption costs may differ across countries, so that low adjustment cost countries adopt new technologies first.

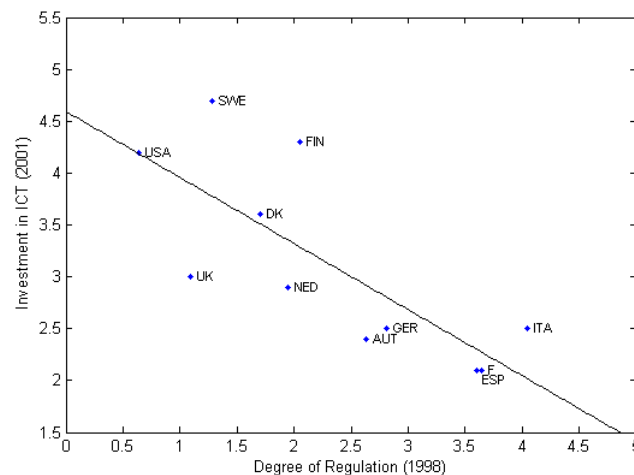


Figure 1.2: Regulation and investment in new technologies, Data is from Nicoletti et al. (1999) and Timmer et al. (2003)

This pattern is also reflected in Figure 1.2, which plots the strictness of an economy's regulatory environment and the degree of technology diffusion, as measured by the ICT investment intensity. The measure of regulation (computed by the OECD) captures the degree of an economy's regulation in product markets, administrative burdens for start-ups and the

degree of state controls. 5 (1) indicates a very strict (loose) regulation. Clearly, countries that have a very strict regulatory environment tend to invest less in new technologies. Gust and Marquez (2002) confirms that differences in regulations are causal for the observed cross country heterogeneity in the adoption of new technologies⁹.

The conjecture that a country's technology adoption behavior is key for understanding its post-1970s labor market performance is supported by comparing the technology/unemployment experience of countries like Sweden, the UK or Denmark with that of countries such as Italy, Spain, France or Germany. Countries belonging to the first group feature a technology adoption behavior similar to the U.S. At the same time, the unemployment rates of those economies are comparable to the U.S. rate. By contrast, economies of the second group seem to adopt new technologies at a substantially slower rate whilst suffering from persistently high unemployment.

1.3 The Model

As an analytical framework I use a vintage technology/vintage human capital model with frictional labor markets. Firms are heterogeneous with regard to the installed production technology. When a new job, i.e. a new production unit, is created it adopts the most advanced technology that is currently available. Each period, firms have the choice of keeping their old technology, upgrading the existing one or destroying the job. Agents are heterogeneous with respect to their human capital endowment. Workers accumulate job-specific skills that are scrapped in the event of a lay-off. The accumulation of specific skills captures the notion of technology learning that increases the productivity of an existing production unit over time. This feature is consistent with empirical results. For instance, Jensen et al. (2001) finds that gains in productivity of an existing production plant resulting from the accumulation of experience are high and significant¹⁰. The specification implies that technologies of different vintages co-exist within productive process of the economy. The use of vintage technology is supported by empirical findings that demonstrate a high persistence of firms' technology¹¹.

⁹More evidence for the negative relationship between a country's technology adoption and the strictness of its regulatory environment is provided by Colecchia and Schreyer (2002) and Jerzmanowski (2006).

¹⁰They find that for a plant that was created in 1967, technology learning accounts for an increase in own productivity of about 57% over the period 1967 – 1992.

¹¹Studies, e.g. by Baily et al. (1992) and Bartelsman and Drymes (1998), suggest that roughly 60% of the firms keep their current level of technology in each period. Moreover, both studies and the findings of Dunne (1994) confirm that plants with relatively poor productivity can restructure their technology and

1.3.1 Vintage Technology and Skills

Time is discrete and denoted by $t \in \{0, 1, 2, \dots\}$. The economy is populated by a continuum of individuals that are either workers or entrepreneurs. Workers are either employed or unemployed. Individuals are infinitely lived but they face a constant probability of death that is given by σ . All individuals are risk neutral and they have no access to savings technologies. An Agent's objective is to maximize expected wealth - which in the absence of a storage technology - is given by the infinite stream of discounted future income.

At each point in time t , there exists a range of sector-neutral technologies denoted by $a_{t,\tau} \in \{a_{t,0}, a_{t,1}, \dots, a_{t,T}\}$ that differ with respect to their date of creation. The vintage of a certain technology is denoted by τ . The leading edge technology is given by $a_{t,0}$ whereas $a_{t,T}$ is the oldest that is still in use. T is determined endogenously and can be interpreted as the critical age at which a technology is scrapped. New technologies arrive at a deterministic rate g . Hence $a_{t+1,0} = ga_{t,0}$, where $a_{t+1,0}$ and $a_{t,0}$ denote the leading edge technologies of tomorrow and today respectively. Newer technologies are, therefore, more productive.

Each employed worker is equipped with a certain stock of human capital, denoted by h . Human capital stock is proportional to the level of specific skills the worker has accumulated on his current job. The accumulation of skills is stochastic. Specific skills, denoted by s , can take on values $s \in S = \{s^0, s^1, \dots, s^I\}$ where $s^0 < \dots < s^I$. s^0 and s^I are, respectively, the lowest and highest potentially attainable skill levels. The transition across skill levels is governed by a Markov process with transition probabilities given by $p(s, s')$. More precisely, $p(s, s')$ denotes the probability that a worker with current skill level s experiences an upgrade of his skills to level s' where $s' \geq s$. Furthermore, I assume $s^0 = 0$, which implies that immediately after the creation of a new production unit or the renovation of an existing one, i.e. when $\tau = 0$ there exists no job/technology specific knowledge. This specification is supported by empirical evidence provided, for example, by Cochran (1960), Garg and Milliman (1961), Rusell(1968) and Pegels (1969). They find that after a change in a firm's production technology, productivity initially drops and then gradually rises. The drop suggests that production knowledge does not apply equally across old and new production technologies and the subsequent increase provides evidence for the existence of learning. The stock of

move up in the relative productivity scale. This can be interpreted as evidence for firm's updating efforts. More explicitly, Dunne (1994) finds that old and young plants appear to use advanced technology at similar frequencies. Given that old plants had installed a different technology when they were created, this implies that at a certain point in time they must have updated their technology.

human capital embodied in a worker is given by

$$h = f^h(s) = (1 + \alpha s) \quad (1.1)$$

where $\alpha > 0$. The functional form of f^h implies that the returns to learning are positive. This is consistent with findings, e.g. by Bahk and Gort (1993) and Jovanovic and Nyarko (1995). Bahk and Gort (1993) uses plant age as a proxy for the vintage of a technology and find that technology learning accounts for an annual increase in output of 1%.

There is a single homogeneous consumption good in the economy that is produced by a continuum of firms. Each firm has a single job that is either vacant or filled with a worker. Firms are heterogeneous with regard to the level of technology specific skills embodied by the employed worker s and the vintage of the implemented technology τ . When a new firm is created it installs the most advanced technology that is available, i.e. $a_{t,0}$. Hence, upon the firm's creation, $\tau = 0$. Firm's output is a function of the installed technology and the worker's stock of human capital. Output is produced according to a constant returns to scale production function given by

$$y_t(\tau, s) = a_{t-\tau,\tau} f^h(s) = a_{t-\tau,\tau} (1 + \alpha s) \quad (1.2)$$

In each period firms have the choice of keeping their old technology, upgrading the existing one by installing the frontier technology or destroying the job. When a firm decides to upgrade, it has to incur a cost χ that is assumed to depend on the technology gap, i.e. the distance of the currently installed technology to the leading edge technology, in the following convex manner

$$\chi_t(\tau) = \mu e^{\xi g^\tau} \quad (1.3)$$

where $\xi > 0$, $\mu > 0$. A firm's relative productivity differential with respect to the technology frontier is denoted by g^τ . We will refer to this differential as a firm's technology gap. The width of this gap is determined by the vintage of a firm's technology τ and the growth rate of the technology frontier. Faster growth at the frontier implies a more rapidly widening technology gap.

In addition to job specific skills, workers are equipped with a set of abilities that enables them to operate the technology they are currently matched with. This particular form of human capital, which we call a worker's production knowledge, does not add anything to

worker's productivity but it is ex-ante required to start operating a certain technology. A worker's production knowledge is characterized by its limited transferability across vintages. Hence, if a worker gets matched with a different (and probably more advanced) technology there is a discrepancy arising between the worker's current production knowledge and that required to operate the new technology. This discrepancy can be best understood by considering the limited transferability as an implicit form of human capital obsolescence. The less a worker's production knowledge can be transferred to a new technology the more obsolete it becomes relative to the production knowledge required for the new technology. Thus, we can understand $\chi_t(\tau)$ as a form of training cost the firm has to incur in order to provide the worker with the necessary skills to enable him to operate the new technology¹².

1.3.2 Unemployment

An existing match can be dissolved for two reasons; exogenous destruction that occurs at the rate ρ or endogenous destruction by firms. A firm destroys its jobs when its production technology is too obsolete and updating was not optimal in the past. Workers that have been laid off are entitled to government sponsored unemployment insurance. This insurance provides the worker with benefit payments during the period of unemployment. Benefit payments, denoted by $b_t(s)$, are a constant fraction ϕ of the average after-tax wage within the respective skill class, i.e. $b_t(s) = \phi(1 - \tau^w) \sum_{\tau} \mu_t(\tau, s) \omega_t(\tau, s)$, where $\mu_t(\tau, s)$ denotes the measure of workers with skills s that are operating vintage τ and $(1 - \tau^w) \omega_t(\tau, s)$ denotes the after-tax wage¹³. Unemployed workers don't receive benefits indefinitely. Instead, each period, an unemployed person that is still entitled to receive benefits faces a constant probability, γ , of losing the entitlement. If that happens, $b_t(s) = 0$ for the remainder of the time spent unemployed.

Therefore, the state vector of each unemployed worker consists of the following three elements: (1) the benefit entitlement, $e \in \{0, 1\}$, (2) the amount of job specific skills accumulated in the previous occupation, s , and (3) the degree of obsolescence of the worker's production knowledge, captured by τ . To better understand the latter suppose that when

¹²We find that the convexity of χ is not critical to our results. We obtain virtually the same results with a linear cost function.

¹³Alternatively, one could make the benefits proportional to the individual's pre-displacement wage, i.e. $b_t(\tau, s) \propto \omega_t(\tau, s)$. We refrain from doing this for two reasons: (a) it would be computationally burdensome given that we would need to keep track of each individuals' last wage income and (b) we think that s and τ are sufficiently correlated across matches so that considering τ would not add any considerable value.

employed in a firm, a worker is attached to a certain technology (with vintage τ) for which he possesses the required production knowledge. This knowledge differs across vintages. Non-zero growth in the leading edge technology and the limited transferability property of the worker's production knowledge imply that a worker's current production knowledge, over time, becomes increasingly obsolete with respect to that required to operate the state-of-the-art technology. The process of human capital obsolescence also applies to unemployed individuals. Consequently, when an unemployed worker gets re-matched with a new firm (one that embodies the leading edge technology) there is a discrepancy between the individual's current production knowledge (linked to vintage τ) and that required to operate the new technology. This discrepancy is captured by τ .

1.3.3 The Labor Market

The labor market is frictional. This means that at each point in time there exists a certain number of open vacancies denoted by v_t and a pool of job-searching individuals, u_t . The total number of unemployed workers is given by $u_t = \sum_{\tau,s,e} u_t(\tau, s, e)$. New matches are determined by a matching function that is homogeneous of degree one, bounded above by $\min\{v_t, u_t\}$ and increasing in both arguments.

$$m_t = m(v_t, u_t) = \bar{m} v_t^d u_t^{1-d} \quad (1.4)$$

where $\bar{m} > 0$ and $d \in [0, 1]$. The probability that a firm meets an unemployed individual with characteristics (s, τ, e) is given by

$$q_t(\tau, s, j) = \frac{m(v_t, u_t)}{v_t} \frac{u_t(\tau, s, j)}{u_t} = m(\theta_t, 1) \frac{u_t(\tau, s, j)}{v_t} \quad (1.5)$$

The last term is implied by the homogeneity assumption on m . Note that $\theta_t = v_t/u_t$ is a measure of labor market tightness. Similarly, let p_t denote the probability that an unemployed worker encounters an open vacancy¹⁴.

$$p_t = m(v_t, u_t) / u_t = m(\theta_t, 1) \quad (1.6)$$

The existence of a matching function in the labor market implies that workers looking for a job trigger a congestion effect on each other. The greater the number of individuals looking for a job the lower is the probability of encountering a vacancy. The same is true for firms with open vacancies. Therefore, a firm's incentive to post a vacancy is strongly affected by the tightness in the labor market.

¹⁴Notice that all heterogeneity is driven by workers as vacancies are all identical, ex-ante.

1.3.4 Government

The government in this economy levies a tax on labor income and redistributes the revenues in the form of unemployment benefits. The government is assumed to run a balanced budget every period. Hence $\tau^w W_t = B_t$, where W is the total wage bill, τ^w is the tax on labor income and B denotes total benefit payments.

1.3.5 Value functions and the firm's updating problem

Given capital-embodied growth there exists a natural trend in the model's key variables. To render the model stationary we divide all growing variables by the common growth factor $a_{t,0}$. To indicate that a variable is stationary we simply remove its time subscript, e.g. $u = u_t/a_{t,0}$.

The Value Functions

All decisions within a firm/worker match - also those concerning matching and technology upgrading - are taken jointly. Regarding the timing of decisions, I assume that a firm and a worker first decide on the technology upgrade and, conditional on the upgrading decision, they then bargain over wages. The value functions for an employed worker and a firm are, respectively, \bar{E} and \bar{J} after the upgrading decision, while before the upgrading decision they are given as E and J . The state variables characterizing a match are the level of skills embodied in the worker, s , and the vintage of the installed technology, τ . Hence, for a given wage rate $\omega(\tau, s)$ we can write \bar{J} as follows:

$$\bar{J}(\tau, s) = y(\tau, s) - \omega(\tau, s) + \beta g(1 - \sigma)(1 - \rho) \sum_{s' \in S} p(s, s') J(\tau + 1, s') \quad (1.7)$$

ρ is the rate of exogenous job destruction, with probability σ the worker dies between two consecutive periods and β is the discount factor. The instantaneous return for a firm is given by output net of wage payments, i.e. $y(\tau, s) - \omega(\tau, s)$. If the match survives to the next period, the age of the installed technology will be $\tau + 1$ and the worker's skill level is $s' \in S$. The value of a job for an employed worker after the upgrading decision is

$$\bar{E}(\tau, s) = (1 - \tau^w)\omega(\tau, s) + \beta g(1 - \sigma) \left\{ (1 - \rho) \sum_{s' \in S} p(s, s') E(\tau + 1, s') + \rho W(\tau + 1, s, 1) \right\} \quad (1.8)$$

This expression is determined by the worker's after-tax wage income given by $(1 - \tau^w)\omega(\tau, s)$ and the discounted future surplus. If the match is hit by an exogenous destruction shock the worker transits to unemployment. Then he is initially entitled to UI benefit claims and the value of unemployment is $W(\tau, s, 1)$.

The updating problem

The cooperative nature of decision making implies that at each point in time a firm/worker pair seeks to maximize the joint surplus of the match. This rule also applies to finding the optimal point in time for a match when to update its production technology. Notice that the production knowledge that an employed worker possesses is outdated compared to that required to operate any newer technology. This is a consequence of constant advancements in the frontier technology and the limited transferability of workers' production knowledge. Both factors render a worker's current production knowledge obsolete with respect to the knowledge required for newer technologies. To overcome this obsolescence, an upgrading firm has to invest in a worker's training, which is costly. The costs associated with a technology upgrade, denoted by $\chi(\tau)$, are assumed to be proportional to the firm's technology gap. The rationale being that the longer a worker is attached to a certain technology, less production knowledge is useable for the new technology and more training will be necessary.

The technology will be renovated when the joint gain from upgrading is positive. This is the case when the following condition holds:

$$\bar{J}(0, 0) + \bar{E}(0, 0) - \chi(\tau) > \bar{J}(\tau, s) + \bar{E}(\tau, s) \quad (1.9)$$

A firm upgrades by adopting the frontier technology, so $\tau = 0$. Upon upgrade, the worker loses his specific skills accumulated with the previous technology, hence $s = 0$ ¹⁵. The upgrading costs are not borne by the firm alone but shared between the firm and the worker. More precisely, total costs, $\chi(\tau)$, are allocated to maximize the surplus i.e.

¹⁵The feature that, upon upgrading, firms jump back to the frontier comes from assumption. However, for the parameterized cost function used here, we would obtain the same result if we let the firm choose the particular technology. This comes from the fact that both the productivity advantage a firm enjoys by adopting any newer technology and the associated costs remain unchanged over time. Hence, to adopt any technology that is less advanced than the frontier would have already been optimal at the time when this technology itself was the frontier technology.

$$\max_{I^E, I^J} [\bar{J}(0, 0) - I^J(\tau, s) - \bar{J}(\tau, s)]^\eta [\bar{E}(0, 0) - I^E(\tau, s) - \bar{E}(\tau, s)]^{1-\eta} \quad (1.10)$$

s.t.

$$I^J(\tau, s) + I^E(\tau, s) = \chi(\tau)$$

I^J and I^E are the adoption costs borne by the firm and the worker respectively and η indicates the firm's weight in the bargain. The outcome of the upgrading problem is a cost sharing rule that solves Equation (1.10) and a policy function, $T(\tau, s)$, which determines, for each pair of states, (τ, s) , the optimal upgrading time, i.e. the maximum age of a technology denoted by τ_s^* ¹⁶.

As previously stated, the bargaining process is sequential. Before the firm and the worker bargain over the wage they decide whether or not to upgrade the production technology. For matters of completeness, we can also state the value function for an employed worker and a firm at the beginning of a period, i.e. before the upgrading decision.

$$J(\tau, s) = \max\{\bar{J}(\tau, s), \bar{J}(0, 0) - I^J(\tau, s), 0\} \quad (1.11)$$

$$E(\tau, s) = \max\{\bar{E}(\tau, s), \bar{E}(0, 0) - I^E(\tau, s), W(\tau, s, 1)\} \quad (1.12)$$

Each period the firm and the worker can also decide to dissolve the match, which will be optimal if the joint surplus of staying in the match - with or without upgrading - is less than zero.

Wage setting

The second step in the intra-match, decision making sequence concerns the wage bargain. Conditional on the outcome of the upgrading decision the firm and the worker engage in a bargaining process in which they choose a wage rate to maximize the joint surplus of the match. The problem they face is given by

$$\max_{\omega(\tau, s)} \bar{J}(\tau, s)^\eta [\bar{E}(\tau, s) - W(\tau, s, 1)]^{1-\eta} \quad (1.13)$$

as before, η indicates the bargaining power of the firm. Optimality implies that

$$(1 - \eta)\bar{J}(\tau, s) = \eta [\bar{E}(\tau, s) - W(\tau, s, 1)] \quad (1.14)$$

¹⁶The cost sharing rule must satisfy $(1 - \eta)[\bar{J}(0, 0) - I^J - \bar{J}(\tau, s)] = \eta[\bar{E}(0, 0) - I^E - \bar{E}(\tau, s)]$.

Using the value function given by Equations (1.7) and (1.8) and the optimality condition (1.14) that has to hold for all pairs of (τ, s) , we can write the wage as

$$(1 - \eta\tau^w)\omega(\tau, s) = (1 - \eta)y(\tau, s) + \eta[W(\tau, s, 1) - \beta g\rho(1 - \sigma)W(\tau', s, 1)] - \eta\beta g(1 - \sigma)(1 - \rho) \sum_{s' \in S} p(s, s')W(\tau', s', 1) \quad (1.15)$$

Match Formation

Cost sharing applies not just to ongoing matches that update their production technology but also to newly formed matches. Unemployed workers suffer from skill obsolescence in the same way as employed workers. To make the production knowledge of a newly hired worker compatible with the installed technology, firms that are newly matched have to invest in workers' training. Similar to the problem above, the costs for a new match are allocated according to

$$\max_{I^E, I^J} [\bar{J}(0, 0) - I^J(\tau, s, e)]^\eta [\bar{E}(0, 0) - I^E(\tau, s, e) - W(\tau, s, e)]^{1-\eta} \quad (1.16)$$

s.t.

$$I^J(\tau, s, e) + I^E(\tau, s, e) = \chi(\tau)$$

Notice that the worker's benefit entitlement affects the degree to which the total training costs are shared. This is intuitive as a worker with UI benefit claims has a higher outside option and a higher opportunity costs of working, compared to a worker with claims. To make such a worker agree on forming the match and leaving unemployment, the firm has to offer a greater share of the total surplus. This is done by letting the worker pay a smaller fraction of the total costs.

The value function of an unemployed worker with benefit entitlement $e \in \{0, 1\}$, that is associated with vintage τ and has accumulated skills s in his last occupation is as follows

$$W(\tau, s, e) = b_e(s) + \beta g(1 - \sigma)p(\theta) [\bar{E}(0, 0) - I^E(\tau, s, e)] + \beta g(1 - \sigma)(1 - p(\theta)) \times [(1 - \gamma_e)W(\tau + 1, s, 1) + \gamma_e W(\tau', s, 0)] \quad (1.17)$$

where the subscript e attached to a variable determines its value conditional on the UI benefit entitlement. The respective variable is 0 for $e = 0$ and equal to its actual value for $e = 1$. E.g. $\gamma_e = \gamma$, ($\gamma_e = 0$) for $e = 1$, ($e = 0$). With probability $p(\theta)$ an unemployed worker encounters a vacancy and becomes employed in the next period. The value of the match

net of training costs is $\bar{E}(0,0) - I^E(\tau, s, e)$. If no match takes place in the current period the worker stays unemployed in the next period as well. In this case - and if still entitled to UI - he loses the benefit claims with probability γ . We have to keep track of a worker's benefit entitlement because, in the case of a match, it will determine a worker's power in the wage bargaining process. A worker not receiving benefits has a lower outside option and has, therefore, a lower reservation wage¹⁷.

In order to fill a vacant job, firms have to actively search for workers by posting vacancies. Denote with V the value of a vacancy and let κ be the cost of keeping the vacancy open. Given free entry, in equilibrium all gains from posting vacancies must be exhausted, hence $V = 0$. In other words, the cost of opening up a vacancy must equal the expected return. The implied zero-profit condition is

$$\kappa = \beta g(1 - \sigma) \sum_{\tau, s, e} q(\tau, s, e) [\bar{J}(0,0) - I^J(\tau, s, e)] \quad (1.18)$$

With probability $q(\tau, s, e)$ the firm meets a worker with characteristics (τ, s, e) , in which case the net surplus of a match is $\bar{J}(0,0) - I^J(\tau, s, e)$. There is no directed search in the model and so the firm can expect to be matched with any worker. This is taken into account by summing over all possible combinations of states (τ, s, e) . Condition (1.18) pins down the degree of labor market tightness in equilibrium.

1.4 Solving the model

The solution to the model is a set of policy functions that characterize the optimal decision behavior of employed workers, firms with a filled job and firms with dormant jobs. Decisions are taken jointly, hence we do not need to separately consider policy functions for employed workers and firms with a filled job. It suffices to focus on a decision rule associated with a job/worker pair. The relevant state variables for an existing match are the vintage of the installed technology, τ , and the level of specific skills, s . For each possible pair of states (τ, s) , the policy function $T(\tau, s)$ determines the optimal decision which is an element in the action space $\{\text{update, not update, destroy the job}\}$ ¹⁸.

The free entry condition that yields the zero (expected) profit condition implicitly determines the vacancy posting behavior of firms with a dormant job. Firms keep posting vacancies as long as the return exceeds the costs, i.e. as long as $V > 0$. Notice that when

¹⁷This is in addition to the effect on the cost sharing rule.

¹⁸An alternative interpretation of $T(\tau, s)$ is that for each skill level $s \in S$ it determines the critical age of the installed technology for which it is optimal to upgrade or to destroy the job.

optimizing, firms and unemployed workers take matching probabilities as given. Therefore, solving the model is reduced to finding a tax rate, τ^w , and labor market tightness, θ , so that the government budget is balanced and ex-post matching probabilities are consistent with agents' ex-ante beliefs.

The algorithm that is constructed to this end is structured in the following way. In an inner loop agents solve their maximization problem taking the tax rate, τ^w , and the tightness, θ , as given. To compute the value functions I use a discrete state space in the τ and the s dimension. The existence of destruction and death shocks ensure the boundedness of the space of vintages τ . The process governing the workers' skill accumulation is a four-state discrete valued Markov process and so the space of possible skill levels consists of four values. Using the resulting policy function $T(\tau, s)$, one can compute the stationary mass functions of unemployed workers $u(\tau, s, e)$ and of existing matches $\mu(\tau, s)$. The matching probabilities (1.5), together with $u(\tau, s, e)$ and the zero profit condition (1.18) is used to update the value of θ leaving the tax rate unchanged.

The inner loop has converged when the value of θ is found that is consistent with agents' prior beliefs. Using the stationary distributions one can compute aggregate variables, including total government expenditures and revenues. In an outer loop, these values are subsequently used to update the guess of τ^w in a way such that the government budget constraint is balanced. The model is solved once the fixed point in τ^w is found. The equilibrium of the model consists of

- a wage schedule $\omega(\tau, s)$ and a firm's policy function $T(\tau, s)$ that maximize the joint surplus of each firm/worker pair (τ, s) ,
- a labor market tightness θ that ensures zero expected profits,
- a tax rate τ^w that balances the government budget and
- distributions of unemployed workers $u(\tau, s, e)$ and matches $\mu(\tau, s)$ that are time invariant

1.5 Parametrization and Calibration

The model period is set equal to half a quarter. In total there are 13 parameters (see Table 1.4) to be set. Seven of them, $(\beta, \sigma, \eta, \alpha, \lambda, d, \xi, g)$, are parameterized, i.e. I use existing micro-evidence to infer about the actual parameter values. The personal discount factor

$\beta = 0.9945$ is chosen so that the implied annualized real interest equals 4.5%. People of working age face a constant probability of dying $\sigma = 0.0025$ so that, on average, they spend 50 years in the labor force. Firms' bargaining weight, η , is equal to 0.5, which is also the elasticity of the matching function with respect to the stock of vacancies. This value is rather standard in the literature on search and matching. The value of $\alpha = 0.2$ is chosen so that the progress ratio i.e. the ratio of peak to initial productivity is equal to 1.2. Jovanovic and Nyarko (1995) reports progress ratios from a dozen empirical studies. Their suggested range is 1.14 – 2.9. Given that 1.2 is a rather conservative choice, we will consider alternative values in Section 1.7. Setting the parameters of the cost function is not an easy task since the

Explanation	Variable	Value	
Discount Factor	β	0.9945	
Probability of dying	σ	0.0025	
Firm's bargaining weight	η	0.5	
Vacancy cost	κ	0.14076	
Parameter of production function	α	0.5	
Parameter of matching function	\bar{m}	0.7702	
Parameter of matching function	d	0.5	
Parameters of cost function	ξ	1	
Growth rate of technology frontier	g	0.025	

		<i>LS</i>	<i>WS</i>
Separation rate	ρ	0.0212	0.0137
Parameter of cost function	μ	0.024	0.024-0.057
Replacement rate	ϕ	0	0.45
Probability of losing benefit entitlement	γ	-	0.0417

Table 1.4: Parameter Values (One period is half a quarter)

empirical literature is rather silent regarding training/updating costs. However, there exists a consensus that training costs are a convex function of the technology gap¹⁹. Therefore, I set $\xi = 1$. The choice of the second parameter in the cost function, μ , will be discussed shortly. The transition probabilities associated with the skill accumulation are set so that intra-firm technology learning lasts, on average, for 10 years. This value is consistent with findings by Bahk and Gort (1993), which reports that intra-firm capital and organizational learning continues for up to 10 years after birth²⁰.

¹⁹However, as stated previously, convexity of the cost function is not critical for our results.

²⁰The lower (upper) bound of the range of attainable skills is set equal $s^0 = 0$ ($s^I = 1$).

Four parameters - $(\kappa, \rho, \bar{m}, \mu)$ - are calibrated in order that their values are chosen such that the steady state generated by the model matches certain features of the U.S. economy for the period before 1975. The pre-1975 steady state of the laissez-faire economy (henceforth *LS*) - which we consider as our benchmark - is calibrated to match the following targets (1) an average duration of unemployment of 11.4 weeks that is consistent with BLS-data for the period 1960-75, (2) an average vacancy duration of 6.5 weeks as reported by van Ours and Ridder (1992), (3) a technology adoption cost to GDP ratio of 2.4% as estimated by Bessen (2002) for 1973 and (4) an unemployment rate of 4%. The parameter values that match these targets are the following. The semi-quarterly vacancy cost $\kappa = 0.14076$ compares to 1.5 months of wage payment, a value that is in line with findings by Bentitola and Bertola (1990) and Felbermayer and Prat (2007). Exogenous layoffs occur with probability $\rho = 0.0212$, i.e. once every 5.9 years. The scale parameter in the matching function is equal to $\bar{m} = 0.7702$ and the curvature parameter of the training cost function is equal to $\mu = 0.024$.

For the welfare-state (henceforth *WS*) there are two more parameters to be set. The replacement rate is set equal to 45%. The OECD reports replacement rates for the early 1970s in Europe lying in the range between 30% (Germany) and 50% (Netherlands). The semi-quarterly probability of losing the benefit entitlement is $\gamma = 0.0417$. It follows that people receive benefits, on average, for 3 years.

For the benchmark case I set the annualized growth rate of embodied technical change equal to 2.5%. Given the differences in the institutions, the same *LS* job destruction rate, ρ , generates a steady state unemployment rate of 5.2% in the *WS* economy. However, average unemployment until the early 70s in Europe was around 3.5%. To account for this fact we follow Hornstein et al. (2007) and recalibrate the separation rate for the *WS* economy²¹. In order to match low European unemployment in the 70s we set the semi-quarterly separation rate equal to $\rho = 0.0137$. A welcome side effect of lowering ρ is that we automatically in-

²¹One might conjecture that introducing layoff taxes - as in Ljungqvist and Sargent (2007) - would be equally successful in resolving the problem. As emphasized by Mortensen and Pissarides (1999), layoff taxes reduce incentives to create jobs and to destroy them. The net effect on labor market tightness (and hence on unemployment) turns out to be ambiguous. In this framework however, a firing tax paid by the firm in the event of an endogenous or an exogenous separation would inevitably raise unemployment. This is due to the fact that, for the current calibration, endogenous job destruction does not exist in equilibrium. The only source of job destruction is exogenous separation. Hence, the channel through which firing taxes could potentially lower unemployment, i.e. via locking workers into their jobs, does not take effect. As a result, layoff taxes would only decrease the surplus of a job for a firm and consequently depress job creation.

crease the average duration of unemployment in the *WS* economy. As reported by Machin and Manning (1999), the duration of unemployment in European countries was already substantially higher relative to the U.S. by the 1970s. We, therefore, believe that our calibration nicely captures the main characteristics of (and differences between) European and U.S. labor markets. The set of parameters and their associated values are compactly summarized in Table (1.4)

1.6 Results

The goal of this paper is to provide an understanding of the linkage between firms' technology adoption behavior, labor market institutions and the labor market performance of an economy. The analysis is motivated by the fact that there are substantial and highly persistent differences in unemployment rates between major European countries and the U.S. and also across European economies. These differences emerged in the late 1970's at a time when there has been a major increase in the arrival rate of new technologies. Given these observations it is clear that we need to evaluate the model along two different dimensions. First, we need to consider a pre-1975 period that is characterized by a low arrival rate of new technologies. The results of this scenario are then contrasted with the results of a post-1975 scenario in which the arrival rate is high. Secondly, in order to mimic the existing technology gap between (and among certain) European countries and the U.S., we need to account for differences in technology updating and is done by considering various different updating cost scenarios. It is not the aim of this paper to explain why Europe has been lagging behind the U.S. in the implementation and usage of new technologies. It takes the observed differences in technology updating as given and seeks to analyze its effects on labor markets outcomes.

1.6.1 The pre-1975 period

Cross-country differences in the frequency of technology upgrading are likely to be the result of different regulatory environments. The evidence presented in Section 1.2 argues that there is substantial heterogeneity in the nature of regulatory environments across European countries. This translates into differences in the underlying adoption-cost structure and causes different rates of technology adoption across countries. In order to account for this we consider different cost scenarios for the *WS* economies. To this end we pick values for the cost parameter μ in the range $0.024 - 0.057$. This range yields average updating costs that are comparable with $4.21 - 13.49$ weeks of average post-update wage payments. Higher costs curb firms' incentives to adopt new technologies and so we observe firms with higher

costs updating their production technology relatively less frequently. Table 1.5 depicts pre-75 outcomes of the calibrated model and contrasts the results of the laissez-faire economy with those of the different cost scenarios of the welfare-state economy.

In the benchmark case, firms upgrade their production technology every 14.44 quarters in the *LS* economy and - depending on the cost structure - every 14.66 – 24.64 quarters in the *WS* economy. The implied technology gap - i.e. the average distance of firms to the leading edge technology ranges from 8.59 – 11.97 quarters.

	L-F, pre 75		Welfare state, pre 75				
Gap to Laissez-Faire, in quarters	0	0.22	2.02	4.37	6.16	8.03	10.2
Average time until update, in quarters	14.44	14.66	16.46	18.81	20.6	22.41	24.64
Average distance to tech. frontier	8.59	8.93	9.6	10.2	10.83	11.44	11.97
Unemployment rate, in %	4.02	3.48	3.73	3.97	4.24	4.58	4.94
Duration of unemployment, in weeks	11.54	14.43	15.51	16.55	17.74	19.2	20.81
% of unemployed w/ spells ≤ 3 months	81.05	69.94	66.43	63.31	60.05	56.43	53.34
% of unemployed w/ spells $[6, 12)$ months	3.46	8.22	10.00	11.65	13.41	15.38	17.03
% of unemployed w/ spells ≥ 12 months	0.13	0.82	1.27	1.81	2.55	3.59	4.9
Equilibrium tax rate, in %	0	1.28	1.37	1.45	1.54	1.65	1.77
L-F: Laissez-Faire Economy ($\phi = 0$), Welfare State Economy ($\phi = 0.45$)							

Table 1.5: Steady states for the period before 1975

The technology choice - labor market linkage

Those differences in firms' technology choice do not leave labor market variables unaffected. For the purpose of understanding the linkages between firms' technology choice and the labor market we first define two concepts that will also prove useful later on. Denote with z_u the average degree of skill obsolescence of unemployed individuals. A worker's skill obsolescence is proportional to his technology gap, so we can write z_u as

$$z_u = \Sigma_{z,s,e} u(z, s, e) z / u$$

where z defines the technology gap in terms of the productivity differential, i.e. $z =$

$1 - e^{-g\tau}$, and u is the total mass of unemployed workers²².

Let c_{tr} denote expected training costs a firm has to incur in the event of a match, which we can write as

$$c_{tr} = \Sigma_{z,s,e} u(z, s, e) \chi(z) / u$$

The maximum age of a technology τ^* can be also expressed in terms of the implied productivity differential, i.e. the maximum productivity gap with respect to the frontier $z^* = 1 - g^{-\tau^*}$. The linkage between the firms' upgrading choice and the labor market is the following. Differences in the updating frequency, τ^* , directly translate into different values of z^* . This is illustrated in rows 1 – 2 in Table 1.6. The longer firms wait to update, the more outdated the technology at the time of the actual replacement. In other words, for less frequent updating the average technology gap of firms z^* increases. Consequently, workers that are attached to these technologies while being employed exhibit a higher degree of skill obsolescence. These workers eventually become unemployed (due to exogenous or endogenous job destruction), making the average degree of skill obsolescence in the pool of unemployed individuals will be higher. This is illustrated in the third row of Table 1.6. Put differently, as z^* rises the distribution of unemployed (across z) shifts to the right and the mass of individuals with relatively more obsolete skills increases.

The degree of skill obsolescence of an unemployed worker determines the amount of training that is necessary in the event of match. Therefore, a higher degree of skill obsolescence in the pool of unemployed implies that firms can expect larger training expenses when being matched with a worker. This is demonstrated in the fourth row of Table 1.6. Depending on the cost regime, expected training costs amount to 6.79% – 17.8% of firm's output. Arguably, larger job creation costs reduce firms' incentives to post vacancies, as illustrated in the fifth row of Table 1.6. Lower job creation leads to a fall in workers' probability to encounter a vacancy. Consequently, the flow out of unemployment is reduced and so causes both the average duration and the level of unemployment to rise. Unemployment rates in the benchmark case range from 3.48% – 4.94%.

Table 1.5 contains results concerning the duration of unemployment. The percentage of jobless workers with unemployment lasting less or equal than 3 months in the laissez-faire

²²Recall that $u(s, j, z)$ is the mass of jobless workers with previous skills s , benefit entitlement e and technology gap z .

economy is around 81.05%, which is close to U.S. data (81.57%, see OECD). The figures for the *WS* outcome are substantially lower (53.34% – 69.94%). This is in line with findings by Machin and Manning (1999). They report that the duration of unemployment in Europe was substantially higher than in the U.S. already in the 1970s. The model also does well in predicting the proportion of long-term unemployed workers in the *LS* economy. The figure produced by the model (3.46%) is close to U.S. data (4.58%). The figure for the *WS* is substantially higher - i.e. it ranges from 8.22% - 17.03% - emphasizing that long-term unemployment was already a severe problem for European economies in the 70s.

	L-S		Welfare State				
Average time until update, in quarters	14.4	14.6	16.5	18.8	20.6	22.4	24.6
Average z^* at update, in %	8.63	8.75	9.77	11.1	12.1	13.1	14.3
z_u , in %	7.39	8.44	8.93	9.35	9.81	10.3	10.7
Average job creation costs, in % output	7.14	6.79	8.81	10.7	12.7	15.2	17.8
Vacancy/employment ratio, $\frac{v}{1-u}$, in %	2.25	1.23	1.15	1.08	1.01	0.93	0.86
Unemployment rate, in %	4.02	3.48	3.73	3.97	4.24	4.58	4.94

Table 1.6: Comparison of cost regimes, pre 75

On the whole, the steady states for the laissez-faire and the welfare-state economy generated by the baseline parametrization can consistently capture the main features of European and U.S. labor markets of the period before 1975.

1.6.2 The post-1975 period

Having analyzed the linkage between firms' technology adoption and labor market outcomes we are now ready to make a next step. Fact *II* presented in Section 1.2 stated that in the late 1970s, there has been a substantial acceleration in capital-embodied technical change. To analyze the effect of this change on labor market outcomes we conduct the following experiment. We increase the growth rate of the technology frontier to $g = 4\%$ to see how this change affects unemployment in economies that differ in their technology updating. As argued previously, a higher rate of arrival of new technologies raises the issue of compatibility problems between vintages. Hence, to overcome the same technology gap adjustment, costs are likely to be higher in a world with rapid technical change.

To account for the increase in adoption costs, for which Bessen (2002) provides empirical evidence, we re-calibrate the cost parameter μ . We set μ to match an adoption cost to output ratio of 6.5%, which concurs with the findings of Bessen (2002) for the U.S. for the period after 1975. The implied value is $\mu = 0.0668$. There is a lack of similar estimates for European countries, hence we determine μ for each of the *WS* scenarios in the following way. The rate of embodied technical change increased uniformly in the U.S. and in Europe, implying we should observe the same compatibility problems in European countries as well. We can expect European firms being confronted with an increase in costs of a similar order of magnitude as those in the U.S. In the calibration we pick the post-75 value of μ so that the relative increase in updating costs for each of the different updating scenarios in the *WS* case exactly matches the increase in costs in the laissez-faire economy. Table 1.7 reports the post-75 steady states of the laissez-faire economy and the different scenarios of the welfare state.

	L-F, post 75		Welfare state, post 75				
Gap to Laissez-Faire, in quarters	0	-1.55	1.26	3.93	6.73	9.61	12.85
Average time until update, in quarters	20.74	19.1	22.01	24.68	27.47	30.35	33.59
Average distance to tech. frontier	9.18	9.15	10.09	10.92	11.76	12.64	13.54
Unemployment rate, in %	6.29	5.48	6.54	7.61	8.84	10.38	12.22
Duration of unemployment, in weeks	18.47	23.22	28.03	32.95	38.82	46.36	55.72
% of unemployed w/ spells ≤ 3 months	58.2	48.42	41.42	36.45	31.25	26.44	22.36
% of unemployed w/ spells [6, 12) months	14.42	19.52	22.54	24.08	24.93	24.38	23.94
% of unemployed w/ spells ≥ 12 months	3.05	7.09	11.87	16.9	22.62	29.72	36.33
Equilibrium tax rate, in %	0	1.94	2.27	2.58	2.93	3.34	3.81
L-F: Laissez-Faire Economy ($\phi = 0$), Welfare State Economy ($\phi = 0.45$)							

Table 1.7: Steady states for the period after 1975

A shock to capital embodied technical change: laissez-faire vs welfare state

An increase in capital-embodied technical change raises unemployment in the laissez-faire economy by about 2.27 percentage points - see Column 1 in Table 1.7. This closely matches the post-75 increase in U.S. unemployment. In the data, we see a jump in U.S. unemployment

in the late 1970s that leads to average unemployment of 6.38% for the period 1975-2000²³. The increase in unemployment is fuelled by an increase in the duration of unemployment which is also in line with the data. In the period 1975-2005, the average duration of unemployment lasted for 15.33 weeks, which is slightly less than the model predicts. From the first columns in Table 1.5 and in Table 1.7 we see that a rise in the arrival rate of new technologies had, on the whole, a rather modest effect on the *LS* labor market.

The effects in the *WS* economy, however, are more diverse. The change in unemployment in response to a shock in capital-embodied growth depends crucially on an economy's ex-ante technology updating frequency. Columns 2 – 7 in Table 1.7 illustrate this pattern. An economy having a technology gap of 4 quarters experiences a rise in unemployment up to 7.61%. For a gap of 6 quarters unemployment increase to 8.84% whereas a gap of 10 quarters pushes unemployment up to 12.22%. These figures broadly match the post-75 unemployment experience of major European welfare state economies with sluggish technology adoption. Leading examples are Germany (7.7%), France (10.55%) or Italy (10.86%). However, as columns 2 – 3 in Table 1.7 demonstrate, unemployment rates in a welfare state with generous UI need not necessarily be high. Economies with sufficiently high upgrading frequencies experience a significantly less pronounced increase in unemployment. In both economies depicted in columns 2 – 3 in Table 1.7 the technology upgrading is virtually the same as in the *LF* economy. Thus, there is practically no technology deficit. The unemployment rates we get for these economies are in the range 5.48% – 6.54%. This is close to the *laissez-faire* figure but, more importantly, it is substantially lower than the rates in welfare states that suffer from sizeable technology gaps. The intuition for this will be provided shortly.

These figures are very much in line with data on countries like Sweden (with an unemployment rate of 5.78%), the Netherlands (6.45%), Denmark (6.67%) or Austria (3.81%). All these economies feature frequent technology upgrading (like the US) and, at the same time, unemployment rates remained at rather moderate levels. Therefore, the key factor that determines the change in unemployment, in response to a shock in capital embodied technical change, is clearly an economy's frequency of technology updates. Generous unemployment benefits do not seem to matter, which is in sharp contrast to the results of Ljungqvist and Sargent (1998).

The non-core role of UI (for the response of unemployment) is supported by the experience

²³Unemployment in the U.S. was mainly above average in the 80s and early 90s. Our model can account for this movement but not for the subsequent decrease that started in the mid-90s.

of classical welfare states such as Sweden, Austria or the Netherlands. All these economies provide generous UI, exhibit no - or just a small - technology gap and at the same time they perform remarkably well in terms of labor market indicators. Overall, this model is able to account for the large variation in unemployment rates we see across European economies. For that, it does not have to rely on institutional differences but rather it stresses the importance of observable technological differences and their linkages to the labor market²⁴.

The effects of capital-embodied technical change on unemployment

The key result of this paper is that, in response to an acceleration of capital-embodied technical change, unemployment rises more than proportionally in economies that feature less frequent technology updating. This pattern is graphed in Figure 1.3. The solid and the dashed lines represent *WS* unemployment rates for the pre-75 and the post-75 era respectively. Each point in the x-dimension represents a particular cost scenario associated with a certain updating frequency. After the simulated shock, the economy jumps to the point on the dashed line that is exactly above the respective pre-75 point. For completeness, the figure also illustrates the response of unemployment in the *LS* economy. Evidently, as the rate of arrival of new technologies accelerates unemployment in the *WS* reaches high levels rather quickly as the technology gap widens.

This section provides intuition behind this result. For this, we make use of the two concepts previously employed, i.e. the average degree of skill obsolescence of unemployed individuals, z_u , and the expected training costs for firms, c_{tr} .

To get a more complete picture of the driving mechanism we also perform the following counterfactual analysis. In a first step we increase only the growth rate of the technology frontier, g , to 4%, and leave the cost parameter μ unchanged. Hence, we disregard the

²⁴It is well known that the persistent increase in unemployment rates in the major European was driven primarily by an increase in the fraction of long-term unemployed. The flows into unemployment were roughly the same in the pre- and the post-1975s era. Outflow rates, however, dropped significantly. As a consequence, the duration of unemployment and the fraction of long-term unemployed rose. This phenomenon is well captured by our model. The results in Table 1.7 show that both indicators experience a significant increase in the post-75 period. The duration of unemployment was generally low in the pre-75 era but experienced a dramatic increase as the rate of arrival of new technologies started to accelerate. The fraction of long-term unemployed (i.e. those with spells ≥ 6 months) in welfare states with sluggish updating ranges from 40.98% to 60.27%. This is broadly consistent with actual data for welfare states like Germany (39%), Spain (51.6%) or France (55.1%). Likewise, the results for economies with frequent updating (26.61% – 34.41%) can broadly match real-world counterparts like Sweden (25.44%) or Austria (37.75%). Also the result for the *LS* economy (i.e. 17.47%) is consistent with the corresponding U.S. figure, (15.53%)

issue of compatibility. In this way we are able to isolate the effects stemming solely from an acceleration in the arrival rate of new technologies. The results are represented by the first rows in Panels I-III in Table 1.8²⁵. In a second step we set g to 4% but also account for compatibility issues by changing the cost parameter μ . We set μ so that the resulting adoption cost to GDP ratio in the *LS* economy equals 4.5%. This is less than in the benchmark scenario for which this ratio is 6.5%. Hence, the problems of compatibility considered are less severe relative to the benchmark case.

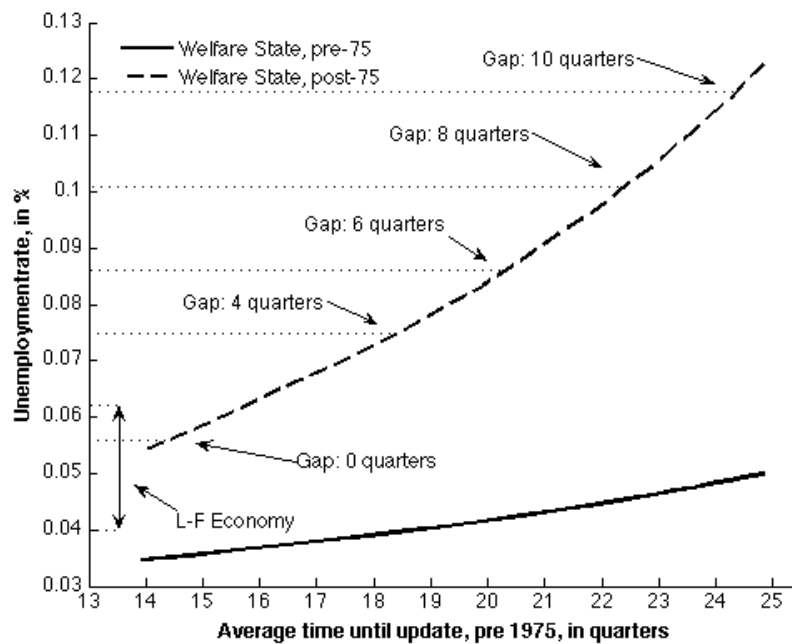


Figure 1.3: Unemployment Rates

The results are given by the respective second rows in Table 1.8. For matters of comparability, we also report the outcome of the benchmark scenario in the respective third rows. The two key elements for understanding the divergence of unemployment rates across economies are (a) the post-75 response of firms' critical gap, i.e. z^* , and (b) the change in the degree of skill obsolescence of unemployed, z_u . We discuss both separately.

From the second column in Table 1.8 we can infer that the critical gap, z^* , increases for all scenarios, even in the case of constant costs²⁶. However, economies that feature sluggish technology updating experience a more than proportional increase in z^* . This is due to the

²⁵Notice that the numbers indicate the change relative to pre-75 outcomes.

²⁶Recall that the critical gap is defined as $z^* = 1 - e^{-g\tau^*}$ where τ^* indicates the maximum age of a technology. For constant costs, T naturally decreases as it is optimal for firms to update earlier. This

fact that as the productivity of new vintages starts to grow faster, implied by a higher g , wages paid within a match drop²⁷. This raises the joint surplus of each existing match²⁸. However, the increase is not the same for all firms.

Cost	Δz^* in %p	Δz_u in %p	ΔJCC , in %	$\Delta \frac{v}{1-u}$, in %	u, in %
Panel I: LS					
=	+3.3	+2.1	+2.4	-9.1	4.4
4.5	+5.3	+2.8	+5.3	-17.5	4.8
6.5	+10.1	+4.8	+14.3	-37.6	6.3
Panel II: $z_{WS}^* - z_{LS}^* = 0$					
=	+3	+2.3	+2.4	-9.7	3.8
4.5	+4.4	+2.9	+5.1	-18.1	4.2
6.5	+8.6	+4.9	+13.7	-37.9	5.5
Panel III: $z_{WS}^* - z_{LS}^* = 10$					
=	+4.8	+3.8	+6.9	-20.7	6.1
4.5	+7.3	+5.2	+14.9	-35.8	7.3
6.5	+14.3	+9.4	+41.3	-62.6	12.2
$z_{WS}^* - z_{LS}^*$ - indicates the technology gap relative to the LS case, %p = change in percentage points, Δz^* = change in firm's critical gap, u = unemployment rate, $\Delta \frac{v}{1-u}$ = change in vacancy/employment ratio, ΔJCC = change in expected costs of job creation, Δz_u = change in average degree of unemployed skill obsolescence					

Table 1.8: Counterfactual Analysis

The longer is the updating horizon, i.e. the higher is τ^* , the longer is the time period for which a firm/worker pair will benefit from lower wages. Consequently, the increase in the joint surplus is higher for matches that update less frequently. A higher surplus postpones the updating date, which implies that the increase in the the critical gap, i.e. $\Delta \tau^*$, will be larger

decrease, however, is outweighed by the increase in the growth rate g so that the overall impact on z^* is negative. As costs increase the maximum age, τ^* , decreases by less or even rises. Hence, in the scenarios that consider also compatibility problems, the critical gap z^* rises by more.

²⁷This is because wages are proportional to a plant's current (relative) productivity, i.e. $\omega(\tau, s) \propto y(\tau, s) = \bar{a}g^{-\tau}(1 + \alpha s)$, see Equation (1.15).

²⁸To be precise, a lower wage increases the firm's surplus, $J(\tau, s)$, but at the same time it reduces the worker's value of the job, $E(\tau, s)$. However, it also reduces the worker's outside option, $W(\tau, s, 1)$, through lowering (a) the unemployment benefits, $b(s) \propto \omega(\tau, s)$, and (b) the value of being re-matched. As such, the overall effect on the joint surplus is positive.

in economies that update less frequently²⁹. Notice that this effect gets more pronounced as we also consider the issue of vintage incompatibility (represented by the middle and far right number in each column). Using the same logic we have developed previously, it is straightforward to understand the remaining figures in Table 1.8. In effect, these results are mostly implied by the response of the critical gap z^* . Recall that z^* determines the degree of skill obsolescence in the pool of unemployed, z_u . The more than proportional rise in z^* for late-updating economies translates into a relatively stronger increase in the degree of workers skill obsolescence, z_u (see column 3 in Table 1.8). This pushes training costs and consequently dampens firms' vacancy posting. As a result, the decline in job creation and therefore, the rise in unemployment is more severe for economies with less frequent technology updating.

The second effect concerns the change in the skill obsolescence of unemployed workers. The degree of obsolescence of workers' production knowledge depends on two factors; (a) the growth rate of the frontier where higher growth implies faster depreciation and (b) the duration of unemployment, as the longer an individual stays out of work then the longer he is exposed to the depreciation process. The first factor implies that the post-75 rise in capital-embodied growth has increased the degree of obsolescence in all economies. However, this negative effect is stronger in economies with sluggish technology upgrading given that workers in those economies face a higher duration of unemployment (see Table 1.7).

1.7 Sensitivity Analysis

1.7.1 The role of unemployment insurance

It is often argued that the generous UI system in Europe is the main reason for high rates of unemployment. Cutting back benefits in our model would, for sufficiently sluggish updating, definitely decreases unemployment. The results in Table 1.7, however, show that even within a welfare state that provides generous UI, low levels of unemployment are achievable. What matters for unemployment is primarily the speed with which firms adopt newly available technologies and the generosity of the public UI system is of second order importance only.

To provide further support for this hypotheses, we run an experiment in which we vary the generosity of the UI system. To this end we pick replacement rates of $\phi \in \{0, 0.25\}$ and

²⁹There is also a second effect that comes through a standard capitalization effect. A higher g increases a firm's surplus by lowering the effective discount rate. Again, firms with a longer updating horizon will benefit more from lower discounting.

compare the outcomes with the benchmark results³⁰. The results are depicted in Table 1.9. For matters of comparability we show just the two extreme updating cases, i.e. those with technology gaps of 0 and 10 quarters. For each of the two cases (organized in Panels I and II) we report the results of each policy regime, i.e. $\phi \in \{0, 0.25, 0.45\}$. The second column contains the change in unemployment in response to the standard pre/post-75 rise in g , i.e. $2.5\% \rightarrow 4\%$. Arguably, the generosity of the UI system (as measured by ϕ) plays a rather limited role for explaining both the post-75 increase in unemployment and the divergence of unemployment rates across OECD countries. For frequent technology updating (Panel I), the response of unemployment is virtually independent of the generosity of unemployment insurance. In the benchmark case, i.e. $\phi = 0.45$, only 0.55 percentage points of the total post-75 increase in unemployment are due to the existence of generous UI. For sluggish updating (Panel II), the effect of benefits is higher. The contribution of UI to the total increase in unemployment equates to 2.37 percentage points. Nevertheless, a large proportion of the increase remains to be explained primarily by firms' technology choice.

ϕ	Δ_u in %p	Duration	z_u	JCC	Δ_{JCC} , in %p
Panel I: $z_{WS}^* - z_{LS}^* = 0$					
0	1.45	17.46	12.99	20.35	-
0.25	1.74	20.35	13.18	20.39	0.47
0.45	2.00	23.22	13.35	20.44	0.92
Panel II: $z_{WS}^* - z_{LS}^* = 10$					
0	4.91	38.07	18.39	56.82	-
0.25	6.61	48.99	19.73	57.51	11.54
0.45	7.28	55.72	20.04	59.02	15.14
Δ_u - change in unemployment rate, %p - change in percentage points, $z_{WS}^* - z_{LS}^*$ - indicates the technology gap relative to the LS case, ϕ = Replacement rate, Duration of unemployment is in weeks, Δ_{JCC} - change in JCC relative to case with $\phi = 0$, JCC = expected costs of job creation, z_u = average degree of unemployed skill obsolescence.					

Table 1.9: The Effect of Unemployment Insurance

The reason why UI has a bigger impact in the case of slow updating is intuitive. For the cases depicted in Panel II, workers stay unemployed for a relatively long period of time

³⁰Note that the case $\phi = 0$ is basically equivalent to a LS economy, but it differs from the previous section in terms of the replacement rate, ρ . In this section, we keep the original parametrization of ϕ for the welfare state, i.e. $\rho = 0.0137$ and only lower ϕ .

implying that they lose a higher proportion of their skills. UI reinforces this effect by prolonging the period of workers' unemployment (see Column 4). Therefore, job creation costs rise by more when updating is sluggish causing a bigger increase in unemployment. However, this should not conceal the fact that the overall importance of UI in explaining the post-75 increase in unemployment is rather limited. This result is in sharp contrast to the results of Ljungqvist and Sargent (1998, 2007), as mentioned previously.

Evidently, there is not much gain from making the UI system more stringent. Reducing the replacement rate to 25% would lower total unemployment by a mere 0.65 – 1.31 percentage points depending on the initial technology gap. More importantly, one should not overlook the dramatic increase in unemployment in an economy that provides just minor unemployment insurance but suffers from large technology gaps. Suppose $\phi = 0.25$ and updating occurs, on average, every 7.5 years. This would generate an unemployment rate of around 9.25% which exemplifies, once again, that very slack technology adoption can have serious consequences for the labor market even in economies that do not provide generous benefits.

1.7.2 Do specific skills matter? The effects on wages and unemployment

The purpose of this section is to analyze to what extent the accumulation of job specific skills affects the equilibrium outcomes. Sluggish technology updating implies that workers operate a certain technology for a relatively long period of time and so accumulate a substantial amount of specific skills, s ³¹. The amount of skills determines an individual's wage and thereby affects his UI claims in the case of unemployment. Consequently, displaced workers with high technology tenure have a valuable outside option and possess strong bargaining power. This reduces firms' surplus in the case of a match and discourages job creation. We will refer to this effect as *skill effect*.

However, there is another effect that potentially counteracts this *skill effect*. When updating is frequent workers operate technologies that are, on average, more productive. Higher productivity translates into higher wages³². We will refer to this effect as *vintage effect*. Which of these two effects dominates is a quantitative question and will depend on the speed and the scope of technology learning.

³¹In the benchmark case of Section 1.6, the average skill level of workers is up to 21% higher in an economy that updates less frequently.

³²Put differently, wages in fast-updating economies are high because workers operate technologies that are, on average, very productive, whereas wages in slow-updating economies are high because workers have accumulated a lot of skills that make them more productive.

To isolate the vintage effect, we first compute the wages that would be paid in an economy without learning. In this way we are able to identify the relative difference in wages that is solely due to the different vintages that are in place. We find that the vintage effect creates a wage differential of 15.46% between in/frequent updating economies³³. We next reintroduce technology learning according to the baseline calibration and we find that the wage differential shrinks to 11.8%. Hence, the skill effect accounts for 3.66 percentage points of the total wage differential. The effect of technology learning on wages under the baseline calibration is rather modest. This is because (1) the scope for technology learning is fairly limited as the progress ratio, the ratio of initial to peak productivity, is 1.2 and (2) the speed of learning is relatively slow³⁴. On average, agents tap the full potential of a technology only after 10 years³⁵. Increasing the speed of learning does not have much effect. If we raise the speed with which skills are accumulated so that the full potential is reached, on average, after 5 (2.5) years the skill effect accounts for 4.4 (4.7) percentage points.

However, the picture changes significantly when we consider higher values of the progress ratio, i.e. when we increase the scope of learning. When we set the progress ratio to 1.5 and consider 10, 5 and 2.5 years of learning we get skill effect amounts of 8.33%, 10.65% and 11.4% respectively. More and faster learning therefore leads to higher wages. One would expect higher wages to cause higher unemployment through raising the workers' outside option³⁶ but is not the case here. By comparing two economies that differ only in the speed or the scope of technology learning we see that unemployment will always be lower in the fast-learning economy. This is intuitive as technology learning implies that workers can raise a plant's productivity at no cost. Consequently, firms' average productivity is higher in economies with learning. In the baseline scenario, average worker productivity exceeds that of an otherwise identical economy without technology learning by 14.3%. As a result, the value of a match for a firm will be higher, which encourages job creation (by 19.18%) and thereby lowers unemployment (by 1.73 percentage points)³⁷. If we increase the speed or the

³³Frequent (infrequent) updating refers to the case of a *WS* economy with a technology gap (relative to the *LS* case) of 0 (10) quarters.

³⁴A progress ratio of 1.2 means that the maximum increase in productivity due to technology learning amounts to 20%. A value that lies in the range 1.14 – 2.9 is suggested by Jovanovic and Nyarko (1995).

³⁵In the benchmark case, this will happen rather rarely given (1) an average firm life of 9.1 years and (2) technology updates that occur (depending on the respective cost scenario) every 4.75 – 8.25 years.

³⁶The rise in the workers' value of unemployment comes through the increase in unemployment benefits that is due to higher wages, $b(s) \propto \omega(\tau, s)$.

³⁷These numbers result from the scenario with sluggish updating. In an economy that updated at the

scope of learning these effects clearly become more pronounced. It may be concluded from this that the more workers can learn about a certain technology, i.e. the more productivity can be raised above its initial level, the higher is the value of a match and hence, the more profitable it is to create new jobs.

1.8 Conclusion

The objective of this paper is to provide a proper understanding of the linkages between an economy's technology adoption behavior, labor market institutions and unemployment. To this end, a labor market matching model was constructed that has been augmented by an endogenous technology choice by firms and a skill accumulation technology for workers. This paper shows that the frequency with which firms in an economy update their production technologies is a key determinant for manner in which the economy's labor market responds to an acceleration in capital-embodied growth. The main result of this paper is that cross-country differences in firms' technology adoption behavior can account for a large part of the observed divergence of unemployment rates across *OECD* economies that are hit by the very same shock (i.e. the acceleration in capital-embodied technical change in the mid 1970s). Unlike previous work in this field, the framework I propose is able to explain both, (a) the divergence of unemployment rates between the major European countries and the U.S. and (b) a large part of the observed variation in unemployment rates across European economies. Furthermore, the results of this paper reject the popular but highly controversial hypotheses that generous unemployment benefits are the main reason for high unemployment in Europe. The analysis shows that even in welfare-state economies with a generous *UI* system, low rates of unemployment are possible if the frequency of firms' technology upgrading is sufficiently high.

The policy implications coming out of this analysis are evident. Rather than thinking about cutting back unemployment benefits, which might create large losses in welfare, policymakers should create conditions that prevent the emergence of a technology deficit. The evidence presented in Section 1.2 hints strongly toward a negative correlation between the strictness of product market regulations and investment in new technologies. Therefore, the removal or the relaxation of burdensome regulatory practices appears to be a natural mea-

frequency of the benchmark *LS* economy the corresponding figures would be as follows: the average productivity and job creation would be, respectively, 13.45% and 7.19% higher and unemployment would be 0.35 percentage points lower.

sure to stimulate rapid diffusion of new technologies by lowering adoption costs. Equally important would be the subsidizing of training of unemployed people, for instance, by providing state-financed unemployment training. This measure would certainly prevent the obsolescence of unemployed workers production knowledge. As a result, it would facilitate the re-integration of individuals into the labor market since it makes it less costly for firms to hire workers³⁸.

³⁸If, however, state-financed training is a substitute for the training provided by firms it could potentially create moral hazard on the firm's side.

Bibliography

- [1] Bahk, B.-H. and Gort, M., 1993, Decomposing Learnings by Doing in New Plants, *Journal of Political Economy*, 101 (4), pp. 561 - 583.
- [2] Baily, M.N., Hulten, C. and Campbell, D., 1992, The Distribution of Productivity in Manufacturing Plants, *Brookings Papers: Microeconomics*.
- [3] Bartelsman, E. and Drymes, P., 1998, Productivity Dynamics: U.S. manufacturing plants 1972-1986, *Journal of Productivity Analysis* 1, 5-33.
- [4] Bessen, J., 2002, Technology Adoption Costs and Productivity Growth: The Transition to Information Technology, *Review of Economic Dynamics*, Vol. 5(2), pp. 443-469.
- [5] Blanchard, O. and Wolfers, J., 2000, The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence, *Economic Journal*, 110.
- [6] Blanchard, O., 2005, European Unemployment: The Evolution of Facts and Ideas, *NBER Working Paper* 11750.
- [7] Bentolila, S. and Bertola, G., 1990, Firing Costs and Labour Demand: How Bad Is Eurosclerosis?, *Review of Economic Studies*, Vol. 57(3), pp. 381-402.
- [8] Cochran, E.B., 1960, New concepts of the learning curve, *J. Industrial Eng.*, 11, pp. 317 - 327.
- [9] Colecchia, A. and Schreyer, P., 2002, The Contribution of Information and Communication Technologies to Economic Growth in Nine OECD countries, *OECD Economic Studies*.
- [10] Cummins, J.G. and Violante, G.L., 2002, Investment-specific technical change in the U.S. (1947-2000): measurement and macroeconomic consequences, *Review of Economic Dynamics*, 5, pp. 243 - 284.

- [11] Daveri, F., 2002, The New Economy in Europe (1992-2001), Gasparini Institute for Economic Research Working Paper No. 213.
- [12] Dunne, T., 1994, Plant Age and Technology Use in U.S. Manufacturing Industries, The Rand Journal of Economics 25(3), 488-499.
- [13] Farber, H.S. The Incidence and Costs of Job Loss: 1982-1991, 1993, Brookings Papers on Economic Activity: Microeconomics.
- [14] Farber, H. S., 1999, Mobility and Stability: The Dynamics of Job Change in Labor Markets, in Ashenfelter, O. and Card, D, (eds), Handbook of Labor Economics, Vol.3, pp. 2439-2483
- [15] Felbermayr, G. and Pratt, J., 2007, Product Market Regulation, Firm Selection and Unemployment, University of Vienna, Working Paper.
- [16] Garg, A. and Milliman, P., 1961, The aircraft progress curve - Modified for design changes, J Industrial Eng., 12, pp. 23-27.
- [17] Gordon, R.J., 1990, The Measurement of Durable Good Prices, NBER Monograph Series, University of Chicago Press.
- [18] Greenwood, J. and Yorukoglu, M., 1997, 1974, Carnegie-Rochester Conference Series on Public Policy 46, pp. 49-95.
- [19] Greenwood, J., Hercowitz, Z. and Krusell, P., 1997, Long-Run Implications of Investment-Specific Technological Change, American Economic Review, 87(3), pp. 342-62.
- [20] Gust, C. J. and Marquez, J., 2002, International Comparisons of Productivity Growth: The Role of Information Technology and Regulatory Practices, FRB International Finance Discussion Paper No. 727.
- [21] Hornstein, A. and Krusell, P., 1996, Can Technology Improvements Cause Productivity Slowdowns?, NBER Macroeconomics Annual 1996, Cambridge MA: MIT Press (1996), 209-59.
- [22] Hornstein, A., Krusell, P and Violante, G., 2007, Technology-Policy Interaction in Frictional Labor-Markets, Review of Economic Studies, 74(4), pp. 1089-1124.

- [23] Jensen, J.B., McGuckin, R.H. and Stiroh, K.J., 2001, The Impact of Vintage and Survival on Productivity: Evidence from Cohorts of U.S. Manufacturing Plants, *The Review of Economics and Statistics*, 83 (2), pp. 323-332
- [24] Jerzmanowski, M., 2007, TFP difference: Appropriate Technology vs. Efficiency, *European Economic Review*, forthcoming.
- [25] Jorgenson, D. W. and Stiroh, K. J., 2000, Raising the Speed Limit: U.S. Economic Growth in the Information Age, *OECD Economics Department Working Papers* 261.
- [26] Jovanovic, B. and Nyarko, Y., 1995, A Bayesian learning model fitted to a variety of empirical learning curves, *Brooking Papers on Economic Activity: Microeconomics*, pp. 247 - 305.
- [27] Ljungqvist, L. and Sargent, T.J., 1998, The European Unemployment Dilemma, *Journal of Political Economy* 106(3).
- [28] Ljungqvist, L. and Sargent, T.J., 2002, The European Employment Experience, *CEPR Discussion Paper No. 3543*.
- [29] Ljungqvist, L. and Sargent, T.J., 2007, Understanding European unemployment with matching and search-island models, *Journal of Monetary Economics*, 54 (8), pp. 2139-2179.
- [30] Machin, S. and Manning, A., 1999, The Causes and Consequences of Longterm Unemployment in Europe, *Handbook of Labor Economics* (ed. Ashenfelter, O. and Card, D.), Vol 3, ch. 47.
- [31] McGuckin, R.H. and van Ark, B., 2001, Making the Most of the Information Age: Productivity and Structural Reform in the New Economy, *Perspectives on a Global Economy*, Report, 1301-01-RR.
- [32] McGuckin, R.H., Spiegelman, M. and van Ark, B., 2005, The Retail Revolution: Can Europe Match the U.S. Productivity Performance?, *Perspectives on a Global Economy*, The Conference Board: New York.
- [33] Marimon, R. and Zilibotti, F., 1999, Unemployment versus Mismatch of Talents: Reconsidering Unemployment Benefits, *Economic Journal*, 109, pp. 266-91.
- [34] Mortensen, D. and Pissarides, C., 1998, Technological Progress, Job Creation and Job Destruction, *Review of Economic Dynamics*, 1(4), pp. 733-53.

- [35] Mortensen, D. and Pissarides, C., 1999, Unemployment Responses to Skill-Biased Technology Shocks: The Role of Labor Market Policy, *Economic Journal*, 109(455), pp. 242-65.
- [36] Nickell, S., Nunziata, L. and Ochel, W., 2005, Unemployment in the OECD since the 1960s. What do we know?, *The Economic Journal*, 115, pp. 1-27.
- [37] Nickell, S., 2003, Labour Market Institutions and Unemployment in OECD Countries, DICE Report 1, CESifo Summer, 13-26.
- [38] Nicoletti G., Scarpetta S. and Boylaud G., 1999, Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation, OECD Working Paper 226.
- [39] Nicoletti G. and Scarpetta S., 2003, Regulation, productivity and growth: OECD evidence, *Economic Policy* 18(36), pp. 9-72.
- [40] Oliner, S. D. and Sichel, D. E., 2000, The resurgence of growth in the late 1990s: is information technology the story?, Federal Reserve Bank of San Francisco.
- [41] Pakko, M. R., 2002, The high-tech investment boom and economic growth in the 1990s: accounting for quality, Federal Reserve Bank of St. Louis, issue Mar., pages 3-18.
- [42] Pegels, C.C., 1969, On startup or learning curves: An expanded view, *AIIE Trans. Eng. Manag.*, 1, pp. 216-223.
- [43] Scarpetta, S., Bassanini, A., Pilat, D. and Schreyer, P., 2000, Economic Growth in the OECD Area: Recent Trends and the Aggregate and Sectoral Level, OECD Economics Department Working Papers, No. 248.
- [44] Schreyer, P., 2000, The Contribution of Information and Communication Technology to Output Growth: A Study of the G7 Countries, STI Working Paper, 2000/2.
- [45] Timmer, M., Ypma, G. and van Ark, B., 2003, IT in the European Union: Driving Productivity Divergence?, Groningen Growth and Development Centre, Research Memorandum GD-67.
- [46] van Ark, B., Inklaar, R. and McGuckin, R., 2002, Groningen Growth and Development Centre, Research Memorandum GD-60.

- [47] van Ours, J. and Ridder, G., 1992, Vacancies and the Recruitment of New Employees, *Journal of Labor Economics*, Vol. 10(2), pp. 138-55.
- [48] Yorukoglu, M., 1998, The Information Technology Productivity Paradox, *Review of Economic Dynamics*, Vol. 1(2), pp. 551-592.

Chapter 2

Informational Frictions and the Life-Cycle Dynamics of Job Mobility

Abstract

This paper studies the life-cycle dynamics of individual job mobility. After entry into the labor market, young individuals typically change jobs very frequently and retain new jobs just for a short period of time. In later stages of their career, workers tend to hold stable jobs and they are significantly more likely to keep a new job than in the early years. This paper argues that the labor market experience that individuals accumulate early in their career affects their job mobility in later stages. We construct a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. The key ingredient is that the imperfection is assumed to be worker-specific and in particular, it is linked to an individual's previous labor market history. We estimate the model by indirect inference on data from the *NLSY 79* and find that it can capture very well the observed life-cycle profile of individual labor market mobility.

JEL Classification: C15, D83, E24, J62

Keywords: Job mobility, Job turnover, Life-cycle, Signaling, Information imperfections, Indirect Inference

2.1 Introduction

Why does individual job mobility systematically differ across age groups? What we observe is that young individuals typically change jobs very frequently and retain each new job just for a short period of time. By contrast, experienced workers tend to hold stable jobs, but more importantly, they are also less likely to separate from a new job than the young. On the aggregate level, this leads to high rates of job turnover and unemployment among the young but comparatively moderate figures for prime-age workers. This pattern is a common characteristic of all labor markets in OECD countries, though it is particularly pronounced in the U.S.

The literature considers the initially high, but declining, turnover for the young to be the outcome of a search process in which workers experiment with jobs in order to find the right match. Accordingly, individuals hold a number of jobs in the first years after labor market entry. Eventually, after a series of failures, one job turns out to be a good match and lasts several years. This explanation is incomplete as it implies that each prime-age worker, who separates from a long-term job and seeks to find a new occupation would, necessarily, have to go through the same wasteful search process as at the beginning of the career. This is clearly not the case. Empirical research has uncovered that the job mobility of *newly employed* workers is declining with age since prime-age workers are significantly less likely to separate from a *new job* than young adults¹. Evidently, mismatch is a phenomenon that is strongly concentrated among young, but appears to be less common among job seekers of higher age cohorts.

The objective of this paper is to contribute to the understanding of the observed pattern of individual job mobility along the life-cycle. To this end, we first present new empirical evidence documenting the life-cycle dynamics of job mobility in the U.S. labor market. Second, in order to account for the observed pattern, we construct a life-cycle model of the labor market. The framework we propose formalizes the hypotheses that the labor market experience individuals accumulate in the initial period after labor market entry affects their job mobility in later stages of the career. Third, to evaluate the empirical content of our

¹Notice that I refer to the reemployment dynamics after a *layoff* and not after a *voluntary separation*. This distinction is important as jobs that follow a voluntary separation are typically more stable than jobs that succeed a layoff. At the same time, we also know that job separations of prime-age workers are more likely to be voluntary than for the young. Due to this "composition effect" in the nature of separations we would automatically observe a difference in the separation hazard across age groups

hypotheses, we estimate the structural parameters of the model using data from the National Longitudinal Survey of Youth 1979 (*NLSY 79*), covering the period between 1979 and 2006. Using the estimated parameters, we perform an empirical validation in which we compare the model's predictions regarding selected individual and aggregate labor market outcomes with their empirical counterparts. In what follows, we provide a brief sketch of each of the three parts of the paper and present some of the main results.

Using data from the *NLSY 79*, we establish a number of empirical facts on individual job mobility in the U.S. We focus on white male labor force participants, mainly for homogeneity reasons. The main facts are:

- The probability of separating from a new job declines significantly with age: 60% of all new jobs among U.S. white males aged 18 – 20 years end within the first year. For workers aged 30 – 34 years, this number is 40% and individuals aged 38 years and over face a 23% likelihood of separating from a new job within a year.
- The survival probability of new jobs drops when we consider a longer retention period: workers aged 18 – 20 years have a 12% (7%) probability of retaining a new job for at least three (five) years, for individuals aged 38 years and over this number is 45% (36%).
- Job mobility is highest when workers start their careers, in later stages job changes become substantially less frequent. At the age of 45 years a typical U.S. white male has held, on average, around 9 full time jobs, 50% of which are held within 5 years of labor market entry and 75% within the first 10 years.
- There is substantial variation in the total number of jobs across individuals. Some individuals will hold very few jobs in their entire career whereas others change jobs frequently. All workers, however, share the common characteristic that the turnover uniformly declines along the career path.
- Unemployment is extremely high for individuals aged 18 – 20 years but it falls sharply, within the age interval 18 – 25 years, from an initial rate of 25% down to roughly 8%. It continues to decline, up to age 35 where it levels out at a rate of 3%. For the remaining part of the working life cycle, that considered here, unemployment stays low and fluctuates in a narrow range between 2% – 3%².

²In the work presented here, we observe each individual up to a maximum age of around 47 years. Thus, we can not capture the unemployment dynamics before retirement.

The evidence presented here hints towards an enormous amount of job mobility and turnover for individuals early in their career. In later stages, job attachments become substantially more durable and job changes less frequent.

To shed light on this issue we construct and estimate a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. We model the imperfection as a noisy signal about the match quality that a firm and worker observe when they first meet. The key ingredient is that the imperfection is assumed to be worker-specific and in particular it is linked to an individual's previous labor market history. Thus, the informativeness of the signal differs across workers as each of them has built up an idiosyncratic labor market history. The intuition is that an individual's labor market history seems likely to convey information about the worker's productive ability in a given job. The more information is available about past jobs the easier it is for the employer to screen the worker and assess their suitability for a particular job. Therefore, young workers are typically associated with less informative signals than experienced individuals, who have already built up an extensive history³.

In the presence of imperfect information, the true quality of a match can turn out to differ from the employer's ex-ante perception. Hence, employers occasionally make mistakes by matching with the wrong workers. The true match quality is an experience good about which agents learn over time, based on observed output. Thereby, mismatches are detected and eventually dissolved. The survival probability of a given match is endogenously determined as it is conditional on the accuracy of the original signal. Given an informative signal, the true quality corresponds to the agents' ex-ante perception, whereas, with noisy information employers are more likely to be wrong. As a result, we observe different patterns of match separation and job mobility across agents that differ from each other in their labor market history and, in particular, in the amount of experience they have accumulated. The life cycle dimension is an important ingredient in this model. It allows us to capture the transition process through which young workers mature (by accumulating labor market experience) and by which they gradually move out of the initial state that is characterized by high job mobility.

³An alternative interpretation is as follows. One might think that the lack of experience causes young individuals to have an inaccurate perception about a particular job, which is expressed by a less informative signal. On the other hand, experienced workers can better assess, ex-ante, the suitability of a given job based on their experience accumulated on past jobs.

We estimate the structural parameters of the model by indirect inference using data from the *NLSY 79*. Each individual in the *NLSY 79* is observed since the point in time when it first enters the labor market until to date (conditional on not dropping out). This allows us to construct the entire labor market history of each individual using the information on the individual's actual labor market transitions. The empirical validation of the model delivers the following results:

- The estimated model can account very well, qualitatively and also quantitatively, for the observed pattern of job mobility along the life cycle. In line with the data, the model predicts that (a) young individuals are substantially more likely to separate from a new job than experienced workers, (b) the hazard rate of separating from a new job is increasing in the length of the retention period and (c) the hazard rate of separating from any given job is declining with tenure, for all age groups. Moreover, the model matches the observed extent of job turnover early in the career as well as the gradual decline in later stages. The implied path for the cumulated number of jobs, at each age, as a fraction of the career total is consistent with observations in the data.
- In the quantitative part, we distinguish between two worker types which can be considered as representing young and experienced workers. The model allows for some of the structural parameters, i.e. those that are governing the signaling process and, thus, the determination of the information imperfection to differ across those two worker types. By how much they actually differ from each other is uncovered by the estimation. We find that the information imperfection, that is integrated into the matching process, is clearly worker-type-specific. Signals are fairly precise for individuals with a great deal of experience, whereas the average labor market entrant is associated with relatively noisy information. This is reflected by a signal to noise ratio that is, the model predicts, roughly 4.5 times higher for experienced workers.
- The fundamental question, however, is for how long a labor market entrant has to stay in the market until the upgrade from high to low noise type is realized? In other words, how much labor market experience does a worker have to accumulate in order to be considered as a low noise type? The model estimates that the cumulated number of years of tenure required for a type change to occur is equal to 5.25.
- As mentioned previously, the quality of a given match is an experience good, about which agents learn over time. Learning is assumed to be "all-or-nothing", such that in each period there is a constant probability that the match quality is either fully revealed or nothing is learned. We estimate this probability and find that it takes, on

average, 1.62 years until the firm (and the worker) learns the true quality of a given match.

- The question that remains is how well the worker-type specificity of informational frictions, which we incorporate, can account for the observed life-cycle pattern of job mobility. To address this issue, we consider a variant of the original model that allows for other factors, which may potentially affect job mobility, to differ across worker types, namely (a) the job quality and (b) the exogenous rate of job destruction. The estimates of the augmented model are then used to disentangle the effect associated with each of the factors and to measure their respective contribution in explaining the observed pattern. The statistic we use in this exercise, to compare the model to the data, is the one-year job retention rate for newly employed workers⁴. We find that the full model can account for 98.5% of the total observed increase in the retention rate over the life cycle. The informational frictions alone can explain a sizable fraction of this increase, 41.8%. Further, allowing the (true but initially unobserved) job quality to systematically differ across worker types adds 45.5% to the explanatory power of the model⁵. The residual 11.2% can be explained by letting the rate of exogenous job destruction to be type-specific.

The question at the core of this paper is: Why are young individuals more likely to separate from a new job than experienced workers? In the work presented here, we focus primarily on the role of informational frictions and we find that they can explain a sizable fraction of the observed patterns. Apart from the information-channel, there are certainly other important factors that shape these patterns. One might think that the young have a higher outside option than prime-age worker or that young adults can often rely on the support of their parents in case of an adverse economic event, such as job loss and unemployment. Typically they, too, have fewer financial obligations regarding child care, the financing of housing, etc. All of these factors are likely to affect their behavior as a higher outside option, generally, makes a job loss less "painful".

⁴The one-year job retention rate measures the probability that a newly employed worker retains her current job for at least one year. We focus on this measure since we believe that it best captures the observed age-specificity of job mobility.

⁵With type-specific job quality, we will observe one type of worker, say the young, to be in jobs that are, on average, worse than the jobs of experienced workers. Reasons for those systematic differences could be factors related to agents' experience, such as (a) agents' search behavior: the young search less directly than the experienced and are, therefore, more likely to end up in a mismatch, or (b) the level of job-specific human capital: the lack of experience may render the jobs of the young less productive.

Another channel concerns agents' search behavior over the life cycle. We may presume that the job search of an individual at the beginning of a career is rather "undirected". Given the lack of labor market experience, young individuals will have a rather inaccurate perception about which job is the right one. Hence, they may consider many jobs as suitable and the likelihood that the one they pick results in a mismatch will be high. Over time, the search behavior becomes more directed as agents learn, on the basis of their performance on previous jobs, for which jobs they are better suited than others. As a result, the range of jobs in which agents search narrows down as the career progresses. Those jobs that are selected, in the end, are likely to be good matches⁶.

The standard job ladder model, as is used by Pissarides (1994), Burdett and Mortensen (1998) and Burdett and Coles (2003), **cannot** be used to explain the decreasing job mobility over the life cycle. More specifically, the job ladder model formalizes the idea that individuals start to work in relatively mediocre jobs but, over time, they climb up the job ladder by quitting their current job and accepting a new one. The process implies that workers select themselves into increasingly better (paid) jobs. Typically, the distribution from which wage offers are drawn has constant support. Consequently, the higher up a worker is in the ladder the less likely he is to receive an even better offer. By contrast, workers in the lower part have a higher probability of receiving a better offer and are, therefore, more likely to separate. Thus, the job ladder model predicts that workers with more labor market experience, who are presumably higher up in the job ladder, face a lower separation hazard than the young.

However, this model exhibits the very counterfactual prediction that all the workers, irrespective of their level of experience, have the same separation hazard **after** they experience a layoff. In the language of the model, a layoff makes workers fall off the ladder and so they will again face the same (high) separation hazard as they did prior to the previous climbing up process. This prediction is clearly contradicted by the data, as it can be observed that experienced workers face a substantially lower hazard rate than the young, **including after a layoff**. The model that is proposed in this paper can account for this feature, as it makes the reemployment dynamics after a separation dependent on the worker's level of labor market experience. The key implication that can be drawn from this is that young

⁶This channel is investigated by Papageorgiou (2009). He proposes a model in which workers learn about their unobserved skills and self-select into the occupation that best matches their abilities. Young workers, who have little experience, are the least informed about their type, they are less likely to choose the occupation in which they perform best and are more likely to switch as a result. This framework can successfully account for the observed decline in occupational mobility with age.

and experienced individuals will face a very different separation hazard when they become reemployed after being laid off.

A sound understanding of individuals' job mobility is key for the design and the implementation of labor market policies. Our results indicate a potentially important path dependency in the working life cycle of individuals. We find that the job situation a worker faces in the later stages of their career is closely linked to the outcomes and the experience accumulated during the early years. Therefore, any labor market policy that affects the working conditions of the young is likely to generate spillovers into the entire working life cycle. The final part of this paper contains further considerations in that direction.

The model presented here draws upon, and extends, the literature on learning in labor market settings. This literature typically assumes that agents initially have limited information about the productivity of a job-worker match but that they learn it over time. The model in Jovanovic (1979) was the first that introduced ex ante uncertainty about job quality into a matching model. In this framework, agents update their prior beliefs in a Bayesian fashion by using the information contained in current period output. Over time, bad matches are detected and eventually dissolve. Pries (2004) proposes an alternative learning mechanism which is "all-or-nothing". Agents observe an initial signal about the match quality but, unlike in Jovanovic (1979), they do not gradually learn the truth. Instead, in each period, the quality of a match is either fully revealed or nothing is learned. The learning process we adopt in this paper is based on Pries (1994). We extend the original approach by assuming that the initial signal, agents observe, is not perfectly informative but noisy. This gives rise to a signal extraction problem, which is key to the informational friction we consider in this paper.

A common assumption in the literature is that firms condition their hiring decision solely upon the initial signal. Other information, and in particular that about workers' characteristics is typically disregarded. Consequently, those models predict that all workers, irrespective of their labor market history, face the same ex ante probability of (a) being hired and (b) of separating from the firm within a given period⁷. This is contradicted by empirical findings. Using a sample of jobs from the *NLSY*, Farber (1994) documents that an individual's current job mobility is strongly positively related to the frequency of job change prior to the start of the job. Gibbons et al. (2005) constructs and estimates a model in which the market and the worker are initially uncertain about some of the worker's skills. In their set-

⁷A notable exception is Lockwood (1991) which assumes that the duration of the preceding spell of unemployment can act as a signal of a worker's productivity.

up, endogenous mobility occurs as labor market participants learn about these unobserved skills. Altonji and Pierret (2001) finds support for the hypothesis that firms statistically discriminate among young workers on the basis of easily observable characteristics such as education. Then, as firms learn about productivity, the importance of the easily observed variables falls but that of hard-to-observe correlates of productivity rises.

The results by Gibbons et al. (2005) and Altonji and Pierret (2001) suggest that the information that accumulates along the career is used by the market to assess the productive ability of a worker. These findings lend strong support to the notion proposed in this paper, namely that the labor market history workers build up along a career is an integral part of firms' information set utilized in hiring decisions. Existing models in the literature typically disregard this dimension and so fail to account for the different patterns of individual job mobility across age cohorts.

Two exceptions are Papageorgiou (2009) and Gorry (2009). Papageorgiou (2009) proposes a model in which workers learn about their unobserved skills along their career and self-select into the occupation that best matches their abilities. In the setup of Gorry (2009), the experience that is accumulated over time aids workers to assess the quality of future job matches. While the underlying mechanism of both is similar to that employed in this paper, they do not explore the implications of learning on individual job mobility. Gorry (2009), for example, focuses on the age related decline in the **aggregate**, but not the conditional, separation rate. That is, Gorry (2009) provides an explanation for why we observe fewer workers of higher age cohorts separating from their jobs but, unlike this paper, does not seek to explain why young and experienced workers with the **same** tenure face different rates of job separation.

The remainder of this paper is organized as follows. In Section 2.2 we present the empirical facts that motivate our analysis. Section 2.3 outlines the structural model that we estimate later on. In Section 2.4 we discuss a simplified version of the full model in order to illustrate the channels through which the considered informational friction affects individual and aggregate outcomes. In Section 2.5 we outline the computational strategy that is adopted to solve the full model. Sections 2.6 and 3.3 present, respectively, the estimation strategy and the data that is used to estimate the structural model. Section 2.8 documents and discusses the estimation results. Section 3.5 closes the main part of the paper with a brief conclusion.

2.2 Empirical Facts

This section uncovers facts that document the variation in individual labor market mobility across age groups. We focus on outcomes that are various measures of job stability and turnover, and unemployment rates.

2.2.1 Individual job stability and turnover

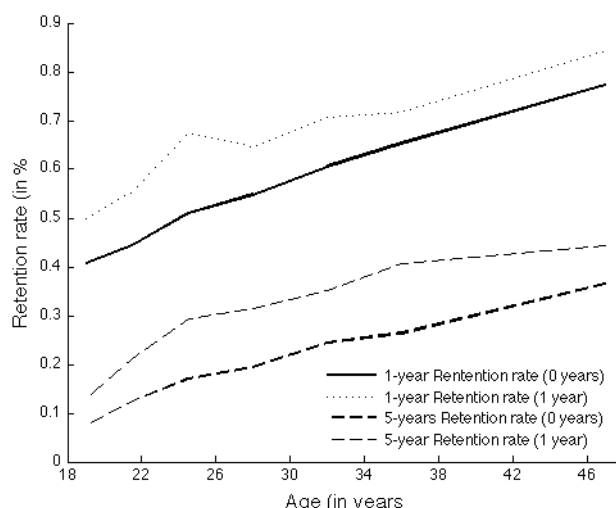
The goal of this section is to establish facts which are meant to illustrate how these measures systematically differ across age cohorts. Farber et al. (1993) and Topel and Ward (1992) find that younger workers have substantially higher rates of job loss than older workers. Job attachments after entry into the labor market are extremely fragile. Topel and Ward (1992) finds that in the U.S., two thirds of all new jobs among young workers end in the first year. However, the probability of job loss declines with experience and jobs become more durable as a career progresses⁸. Their results imply that a worker with 12 years of experience is about half as likely to leave a new job as a new entrant. Hall (1982) arrives at a similar conclusion and estimates that the 5-years job retention probability for male workers aged 15 – 24 is only 3.93%, whereas the rate for more experienced workers is substantially higher. For workers aged 35 – 39 (45 – 49), this probability is equal to 16.0% (20.0%).

The job retention probability is a useful conceptual measure of job stability and refers to the probability that a worker with a given age and tenure will retain his current job for a certain number of years. Figure (2.1) and Table (2.1) report job retention probabilities calculated from *NLSY* 79 data. The sample considered here consists of U.S. white male workers. Individuals attending school or serving in the army are excluded. Moreover, we consider only full time jobs, i.e. jobs with ≥ 30 weekly hours. The panel dimension of the *NLSY* 79 allows us to construct the entire labor market history of each individual in the sample. Therefore, the reported retention rates are based on the actual labor market transitions of each observed individual. In particular, the rates are calculated from the number of workers in a given age-tenure category who move on to higher age-tenure categories. For example, to compute the probability that a worker aged a years, who has been on the job for k years and who will remain on the job for s more years, we use

$$\frac{\text{Number of workers aged } a + s \text{ years with } k + s \text{ years of tenure}}{\text{Number of workers aged } a \text{ years with } k \text{ years of tenure}} \quad (2.1)$$

⁸Notice that the probability of job loss in this context refers to the probability of losing a *new* job.

Figure (2.1) depicts the $s \in \{1, 5\}$ -year job retention rates for workers aged $a \in \{18, 46\}$ years with tenure $k = \{0, 1\}$, i.e. considering newly employed workers and those with one year of tenure.



Comparison of the 1-year and 5-years job retention rates for workers aged 18-46 years with 0 years and 1 year of tenure. Source: Own calculations using data from the *NLSY 79*

Figure 2.1

Three things are worth noting: (1) The job retention rate for workers with one year of tenure is always higher than the corresponding rate for newly employed workers. In other words, tenure substantially increases the likelihood of retaining the current job. This is not surprising as it is well known that the hazard rate of separating from an employer declines with tenure. This relation is also reflected here. (2) Also as expected, the 5-years retention rate is always lower than the 1-year rate. The longer the retention period the more likely it is that the job gets destroyed. (3) Most importantly, all reported job retention rates exhibit a strong increase with age. Evidently, young cohorts are much more likely to leave a new job than experienced workers. This pattern holds for all tenure classes and for all retention horizons.

The numbers underlying Figure (2.1) are reported in Table (2.1). The numbers in brackets are the corresponding retention rates for workers with one year of tenure. A worker aged 18 – 20 starting a new job has a 40% probability of retaining this job for at least 1 year. For an individual 30 – 34 years old, this probability is already 60% and a worker with roughly

20 years of labor market experience has a likelihood of 77% of retaining a new job for one year or longer.

Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.41 (0.49)	0.51 (0.67)	0.60 (0.71)	0.77 (0.84)
2 years	0.21 (0.31)	0.33 (0.51)	0.42 (0.57)	0.61 (0.58)
3 years	0.12 (0.21)	0.26 (0.42)	0.33 (0.43)	0.45 (0.56)
4 years	0.09 (0.15)	0.21 (0.33)	0.27 (0.39)	0.41 (0.47)
5 years	0.07 (0.13)	0.17 (0.29)	0.24 (0.35)	0.36 (0.44)

Job retention probabilities for workers with a given age (at the time of recruitment). Rows represent different job retention periods. Numbers in (without) brackets are the rates for workers with 52-103 (1-51) weeks of tenure. Source: Own calculations using data from the *NLSY 79*

Table 2.1

The retention probability drops significantly when we consider a longer retention horizon. For young age cohorts the probability of keeping a new job for at least three (five) years drops to 12% (7%) whereas the corresponding rate for experienced workers is 45% (36%). The numbers reported here are slightly higher than those found by Hall (1982). This is not surprising since the sample of workers considered here consists just of white males whereas Hall (1982) includes both females and blacks; two groups that are known to demonstrate less stable job attachments. To sum up, young workers when finding a job face a much higher probability of losing this job within a relatively short period of time.

Evidently, mismatch is a phenomenon that is strongly concentrated among young cohorts. High rates of job loss for the young imply a lot of job turnover early in the career, a phenomenon which is also known as job-shopping. According to Topel and Ward (1992), a typical US male worker will hold seven full-time jobs, which is about two thirds of his career total, during the first 10 years in the labor market⁹. Own calculations, using the aforementioned *NLSY 79* data, confirm these results. The first row in Table (2.2) depicts the mean number of (full-time) jobs for individuals within a given age. Standard deviations are given in parentheses.

⁹The sample in Topel and Ward (1992) is taken from the Longitudinal-Employer-Employee-Data (LEED), which is based on individuals' Social Security earnings records.

Up to age 45, a typical US white male worker has held, on average, 8 – 9 full time jobs. Notice that the "accumulation" of jobs over the life cycle is far from being uniform. The lower panel of Table (2.2) reports the number of jobs held, until a given age, as a fraction of the career total. 50% of the total jobs are held within 5 years of labor market entry and 75% of the total number is held within the first 10 years (assuming that labor market entry occurs at age 18). This hints at an enormous degree of job turnover among young individuals. In the later stages of a career, job changes become substantially less frequent. Table (2.2) illustrates that the increase in the number of jobs flattens out at roughly 30 years of age.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
Total number of lifetime jobs							
	2.58	3.84	5.18	6.61	7.42	8.36	9.18
	(1.68)	(2.36)	(3.01)	(3.78)	(4.31)	(4.88)	(5.36)
Number of jobs as a fraction of total number							
	0.332	0.484	0.629	0.746	0.806	0.882	0.963
	(0.211)	(0.236)	(0.232)	(0.211)	(0.201)	(0.165)	(0.084)
Upper panel: Mean number of (full-time) jobs for individuals until a given age. Lower panel: Mean number of jobs held until a given age, as a fraction of the career total. Standard deviations are in parentheses. Source: Own calculations using data from the <i>NLSY 79</i>							

Table 2.2

Notice that there is substantial variation in the total number of jobs across individuals. The standard deviation associated with the average number of jobs is strikingly high. This suggests a large degree of heterogeneity regarding job changes. Some individuals will hold very few jobs in their entire career whereas others change jobs frequently. All workers, however, share the common characteristic that the turnover uniformly declines along the career path. This can be seen from the standard deviations reported in the lower Panel of Table (2.2). The turnover pattern we have identified does apply not solely to the U.S. labor market but it is a common characteristic of advanced OECD labor markets. Evidence for other countries exists for the UK and for Germany. Booth, Francesconi and Garcia-Serrano (1997) reports that half of the average five UK job changes occur in the first ten years of labor market participation. For Germany, Winkelmann (1997) finds that almost half of the average (four) German job changes fall in this period.

High rates of job turnover for young cohorts imply that the average tenure of each job is low. Table (2.3) depicts the average tenure of currently employed workers and the average duration of matches that ended.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
Average tenure							
	0.87 (0.91)	1.46 (1.38)	2.13 (2.06)	3.11 (2.91)	4.43 (3.95)	5.85 (5.09)	7.91 (6.77)
Average duration of dissolved matches							
	0.494 (0.58)	0.751 (0.92)	1.094 (1.33)	1.429 (1.79)	1.931 (2.47)	2.681 (3.35)	3.665 (4.52)
Upper panel: Average tenure of currently employed workers with a given age. Lower panel: Average duration of matches that ended for workers with a given age. Standard deviations are in parentheses. Source: Own calculations using data from the <i>NLSY</i> 79							

Table 2.3

The average tenure increases in age as expected and at the same time, the average duration of dissolved matches is roughly half of the duration of existing matches. This is interesting, especially for prime-age cohorts, as it suggests a coexistence of stable life-time jobs and high turnover. To shed more light on this issue, we calculate the fraction of workers within a given tenure class. The results are reported in Table (2.4).

For young cohorts the majority of the workers are in short-term jobs. This is implied by (a) the high job turnover of young workers and (b) the limited amount of time that has passed since labor market entry. When we consider later stages of the working life cycle, more and more individuals can be found holding medium and long-term jobs, though the fraction of short-term jobs among prime-age workers is still very high¹⁰. For the cohort aged 38 and higher, one finds that there exists a strong dualism in the labor market. A substantial fraction of workers are in very stable jobs but at the same time a high proportion of individuals are subject to high turnover, holding many jobs for a short period of time only. 33% of workers aged 38 years or older are in jobs lasting for 10 years or more. At the same time, almost one quarter of workers in the same age cohort have less than 2 years of tenure.

¹⁰The term "prime-age" refers to workers aged 24 – 54.

The general pattern, regarding job mobility over the life cycle, that emerges is that young workers are subject to high turnover and generally face a high probability of job loss. In contrast, experienced workers are less affected by mismatch and usually end up in more stable jobs. After labor market entry, young individuals typically hold a number of very brief jobs in the first few years. Eventually one job turns out to be a good match and lasts several years.

Tenure / Age	[18-20)	[23-26)	[30-34)	[38-
0 – 2	0.8952	0.6062	0.3715	0.2394
2 – 5	0.1011	0.2733	0.2645	0.2134
5 – 10	0.0036	0.1189	0.2446	0.2156
> 10	0	0.0017	0.1196	0.3316

Fraction of workers of a given age within a certain tenure class. Source: Own calculations using data from the *NLSY* 79.

Table 2.4

2.2.2 Unemployment

Unemployment and job turnover are naturally linked to each other since a job separation is often followed by a period of unemployment. The findings of the previous section therefore suggest that we should also observe a strong age dependency for unemployment. Panel (a) in Figure (2.2) shows the average unemployment rates for young and prime-age male workers for the period 1985 – 2004 for selected OECD countries¹¹. Each of the points represents the combination of young-age and prime-age unemployment rates for a particular country. If a point lies above the 45° degree line, then the unemployment rate of the young exceeds that of prime-age workers and vice versa. For all the countries that are considered here, we observe a large discrepancy in unemployment rates between both age groups. Low rates of unemployment for prime-age workers typically coexist with high rates for young cohorts. In addition, there is substantial cross-country variation in the difference between both rates. In

¹¹We focus on male workers only for matters of comparability. Arguably, some of the variation across age groups is likely to be driven by differences in labor force participation between male and female workers. Hence, putting both groups together would distort the picture.

absolute terms, the difference ranges from a minimum of 1.97 percentage points in Germany to 20.71 percentage points in Italy. Most of the countries can be found in the range of 7 – 13 percentage points.

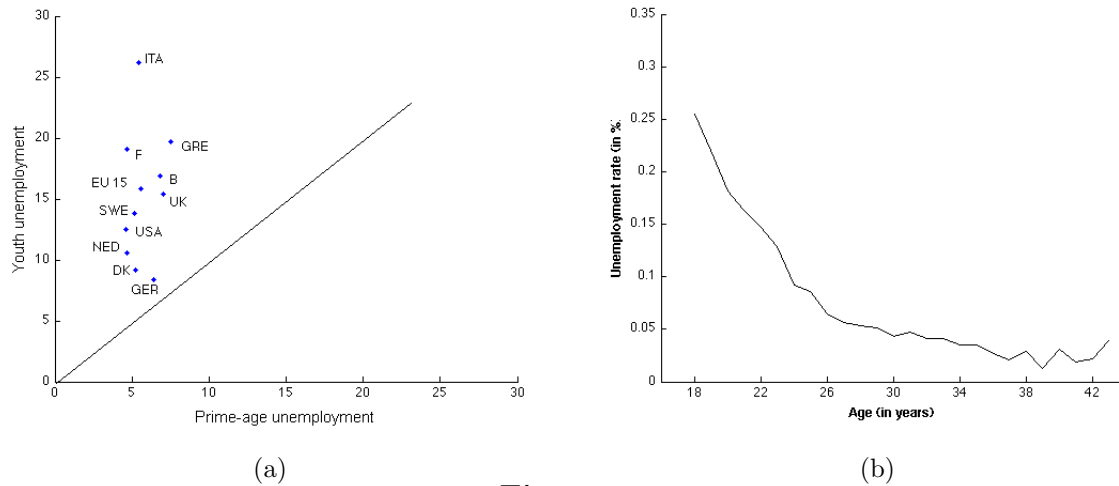


Figure 2.2

Panel (a): Cross-country comparison of the relation between the unemployment rates for young workers, aged 18-24 years, and prime-age workers aged 24-54 years. Source: OECD Labor Force Statistics. Panel (b): Unemployment rate for U.S. white male worker aged 18-43 years, Source: Own calculations using data from the *NLSY 79*.

Panel (b) in Figure (2.2) provides a more detailed view for the U.S. labor market. In particular, it draws the unemployment rate for U.S. white males workers belonging to different age cohorts. Unemployment is extremely high for individuals aged 18 – 20 years but it falls sharply, within the age interval 18 – 25 years, from an initial rate of 25%, down to roughly 8%. It continues to decline, up to age 35 where it levels out at a rate of 3%. For the remaining part of the working life cycle, that considered here, unemployment stays low and fluctuates in a narrow range between 2% – 3%¹².

2.3 The Model

In this section, we construct a life-cycle model of the labor market that is used to gather insights on the pattern of individual job mobility identified above. The building blocks of the model are (a) firms and deterministically aging workers, matching in the labor market, (b)

¹²Notice that due to the limited time horizon of the *NLSY 79*, we do not observe individuals until retirement.

an informational friction that distorts the hiring process and introduces uncertainty about the quality of a job/worker match, (c) employers, learning about the match quality over time and (d) workers, building up an idiosyncratic labor market history which affects the informational friction. In what follows, we discuss each of the elements in turn.

2.3.1 Workers

Time is discrete and denoted by $t = 0, 1, 2, \dots$. The economy is populated by a continuum of risk-neutral individuals that are facing a finite life cycle. Let $k \in \mathcal{K} = \{0, 1, \dots, K\}$ denote an individual's age. The aging process is deterministic, i.e. an individual aged k at time t will be of age $k + 1$ in period $t + 1$. Individuals face a certain probability of dying that is age dependent. Let $\rho_k \in [0, 1]$ denote the conditional survival probability from age k to age $k + 1$. Individuals live up to a maximum of K periods. At age K agents die with probability one, i.e. $\rho_K = 0$.

Worker, whether employed or unemployed, are characterized by their "type" denoted by i , where $i \in \mathcal{I} = \{1, 2, \dots, \bar{I}\}$. The type is meant to capture all the "ability-related information" that could be of value to an employer in order to, ex-ante, screen and categorize a worker and to assess his productive ability for a given job. We can think of this information as being based on factors such as curricula vitae, recommendations, personal interviews, test scores and the like. If a worker is of a high i -type, then more specific information is available upon which an employer can condition his hiring decision. Instead, for a low i -type worker, there is little information available. Arguably, with little information about an applicant, it will be hard for the employer to draw inferences on the worker's productivity and the quality of a potential match. In other words, when facing low i -type workers it will be hard for an employer to distinguish stars from lemons. The informational friction that we consider here is based on exactly this logic, but we will be more specific on this issue shortly.

Clearly, the amount of information that is available is related to workers' labor market experience. The longer an individual is in the market, the more information will be available on the basis of which employers can judge a worker's qualification for a given job. Therefore, i does not remain fixed over an individual's working life cycle but it changes over time. We make the following assumptions about the law of motion of i :

- At the moment of labor market entry, the worker's type is equal to the lowest possible value, i.e. when $k = 0$ we have that $i = 1$.

- The law of motion of i is stochastic and is governed by a discrete valued Markov process with certain transition probabilities. These probabilities depend on the respective labor market state of the worker. Let $\mu_{i|i'}^s$ denote the probability that a worker with labor market state s experiences a change of his type from i to i' within the current period. The set of labor market states we consider include $s = \{e\text{-employed, } u\text{-unemployed, } b\text{-match break-up}\}$.

Workers do not have access to a savings technology and so consume their entire income every period. Preferences over consumption are assumed to be representable by a standard time separable utility function of the form $E \{ \sum_{t=1}^{\infty} \beta^{t-1} u(c_t) \}$ where β is the time discount factor and c_t represents an individual's level of consumption. Expectations are taken with respect to the stochastic processes that govern mortality and the matching and signaling processes in the labor market.

2.3.2 Firms

A firm is assumed to consist of just one job. A firm's job can either be vacant or filled. Notice that there is no ex-ante heterogeneity on the firm's side, i.e. before being matched with a worker, firms are all homogenous. In order to produce output, a firm has to be matched with a worker. A firm/worker match can be of good or bad quality. Good matches produce y^g whereas bad matches produce y^b , with $y^g > y^b$. The match quality is an experience good, meaning that it is unknown to the firm and the worker when they first meet. The probability that a match between a type i worker and a firm is good quality is given by π , with $\pi \sim N(\alpha, \sigma_\pi^2)$ where $\alpha, \sigma_\pi^2 \geq 0$ and π is unknown. Instead, agents (the firm and the worker) observe a noisy signal, about π , which is denoted by γ . More precisely, the signal γ is assumed to consist of the true value π and a noise component ϵ_i , i.e. $\gamma \equiv \pi + \epsilon_i$. The noise component is assumed to be worker specific. In particular, we make the following assumptions about ϵ

- $\epsilon_i \sim N(0, \sigma_{\epsilon,i}^2)$ where $\sigma_{\epsilon,i}^2 \geq 0 \forall i$
- $\sigma_{\epsilon,i'}^2 \leq \sigma_{\epsilon,i}^2$ for $i' \geq i$

The latter condition implies that the signal of low i -type workers will contain, on average, more noise than the signal of high i -type workers. From the assumptions above, it follows that $\gamma \sim N(\alpha, \sigma_\pi^2 + \sigma_{\epsilon,i}^2)$. The firm and the worker form beliefs about π on the basis of the observed signal γ . Let $\hat{\pi}$ denote the expected probability that the match is good, conditional

on having observed the signal γ . Given the normality assumptions about π and ϵ we can write $\hat{\pi}$ as

$$\hat{\pi} \equiv E(\pi|\gamma, i) = (1 - \eta_i)\alpha + \eta_i\gamma \quad (2.2)$$

where $\eta_i = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_{\epsilon,i}^2}$. It follows that $\hat{\pi} \sim N(\alpha, s_i^2)$ where $s_i^2 = \sigma_\pi^2 \eta_i$. Notice that the worker's type is key for the formation of beliefs. The higher the type of a worker the smaller the variance of the noise component $\sigma_{\epsilon,i}^2$, and hence, the signal will contain more information about the true match quality. Therefore, more weight will be put on the signal γ . On the other hand, if $\sigma_{\epsilon,i}^2$ is high, meaning that the signal will be rather uninformative, more weight will be put on the mean α . In the extreme case of a totally uninformative signal - $\lim_{\sigma_{\epsilon,i}^2 \rightarrow \infty} \hat{\pi} = \alpha$.

2.3.3 Match Formation

When a type i worker meets a firm with an unfilled job, both parties observe the signal γ and form beliefs about π . Those beliefs are expressed by the expected conditional probability $\hat{\pi}$. Conditional on the value of $\hat{\pi}$, both parties decide whether or not to form a match. When the match is not formed both continue their search and when the match is formed, the firm starts producing output in the subsequent period. The state of a firm with a filled job is given by the triplet $(k, i, \hat{\pi})$ consisting of the age (k), the type (i) of the employed worker and the common beliefs, $\hat{\pi}$.

The output of a match depends on the match quality. In matches of unknown quality, the firm and the worker don't observe the true output directly but rather $y = \bar{y} + e$, where $\bar{y} \in \{y^b, y^g\}$. If output was perfectly observable, all the uncertainty about the match quality would vanish immediately after the match formation. We assume that the realizations of the output noise, e , are drawn from a uniform distribution with support $[-\mu, \mu]$. At first, the quality of a match is unknown. However, workers and firms learn about it over time on the basis of observed output. Given the assumptions on e , the learning process will be "all-or-nothing". This approach of modeling the learning of match quality was first proposed by Pries and Rogerson (2005).

The name "all-or-nothing" refers to the manner in which new information, that is generated within a period, is processed by the agents. In particular, if the realization of current output is such that $y > y^b + \mu$ ($y < y^g - \mu$) the agents can infer that $\bar{y} = y^g$ ($\bar{y} = y^b$). This implies that **all** the uncertainty about the match quality is removed at once. On the other

hand if $y^g - \mu \leq y \leq y^b + \mu$ the agents have learnt **nothing** about the truth. Therefore, the probability that the match quality is revealed within a period is given by $\varphi = \frac{y^g - y^b}{2\mu}$, which is constant over time and independent of tenure. When a match is found to be bad it dissolves. On the other hand, good matches remain intact until either the worker dies or the match is hit by an exogenous separation shock.

2.3.4 Value Functions

Unemployed Workers

Workers can be either employed or unemployed. There is no labor force participation or retirement choice in the model. Unemployed workers encounter firms with open vacancies at an exogenous, Poisson rate given by $\tilde{\lambda}$. Let $\Gamma^u(k, i)$ denote the value of unemployment to an individual with age k and type i . We can write $\Gamma^u(k, i)$ as follows

$$\Gamma^u(k, i) = b + \beta \rho_k \tilde{\lambda} \left(\int_{\hat{\pi} \in [0,1]} \Theta_u(k', i, \hat{\pi}) dG^i(\hat{\pi}) \right) + \beta \rho_k (1 - \tilde{\lambda}) \sum_{i' \in I} \mu_{i|i'}^u \Gamma^u(k', i') \quad (2.3)$$

where $\beta \in [0, 1]$ is the personal discount factor and ρ_k denotes the probability that the worker survives into the next period. The value of unemployment consists of three parts. The first part, given by $b \geq 0$, is the flow value of unemployment. We can think of b as representing, for instance, the value of leisure, unemployment benefits and the like.

The second part is the expected discounted value of encountering a firm with an open vacancy. Upon meeting a firm, a signal γ is drawn on the basis of which agents form beliefs $\hat{\pi}$. There is a one-to-one mapping between γ and $\hat{\pi}$. Therefore, we can omit the signal in the value function and work directly with the corresponding beliefs $\hat{\pi}$. Let $G^i(\hat{\pi})$ denote the cumulative probability distribution function associated with $\hat{\pi}$. Notice that the distribution of beliefs is worker-specific as the variance of the noise component $\sigma_{\epsilon, i}^2$ is a function of the worker's type.

The realized value of $\hat{\pi}$ will be crucial for whether or not a match is created. In equilibrium there will be an endogenous threshold for $\hat{\pi}$, above (below) which agents find it optimal (not optimal) to form the match. We will be more explicit about the determination of this threshold shortly. If the draw is sufficiently good then the match is formed, i.e. the worker agrees to stay with the firm and vice versa. The benefit of accepting a job offer, i.e. the value

of employment, is given by $\Gamma^e(k', i, \hat{\pi})$. If the worker finds it optimal to reject the job offer he continues searching. In this case, the expected value of unemployment for next period is $\sum_{i' \in I} \mu_{i|i'}^u \Gamma^u(k', i')$. Notice that next period's value takes account the possibility of a change in the worker's type i . Thus, we can write the maximum value of meeting a firm as

$$\Theta_u(k', i, \hat{\pi}) = \max \left\{ \Gamma^e(k', i, \hat{\pi}), \sum_{i' \in I} \mu_{i|i'}^u \Gamma^u(k', i') \right\} \quad (2.4)$$

The expected value of a meeting is obtained by integrating (2.4) over all possible realizations of the beliefs $\hat{\pi}$. The third part of expression (2.3) is the expected discounted value of remaining unemployed. This value is realized with probability $(1 - \tilde{\lambda})$, which is the likelihood that an unemployed worker does not encounter a firm with a vacant job.

Firms

The state of an existing match is given by the triplet $(k, i, \hat{\pi})$. Let $J(k, i, \hat{\pi})$ denote the value of a job to a firm that employs a type i / age k worker, who has beliefs $\hat{\pi}$. We can write J as follows¹³

$$J(k, i, \hat{\pi}) = \max \left\{ E(y|\hat{\pi}) - w(k, i, \hat{\pi}) + \beta(1 - \sigma)\rho_k J^+(k, i, \hat{\pi}), 0 \right\} \quad (2.5)$$

where

$$J^+(k, i, \hat{\pi}) = \sum_{i' \in I} \mu_{i|i'}^e [(1 - \varphi)J(k', i', \hat{\pi}') + \varphi\hat{\pi}J(k', i', 1)]$$

A firm's outside option is always the value of an unfilled vacancy, which in equilibrium is equal to zero. Therefore, the max operator represents the firm's optimal choice with respect to current period employment. The value of a job to a firm consists of the two parts: the instantaneous return and the continuation value. The latter is denoted by $J^+(k, i, \hat{\pi})$, whilst the instantaneous return is given by the difference between expected output, $E(y|\hat{\pi}) = \hat{\pi}y^g + (1 - \hat{\pi})y^b$, and the wage, w , that is paid to the worker. Wages are

¹³It is important to note that due to the all-or-nothing learning, agents' beliefs, $\hat{\pi}$, stay the same throughout the period in which the quality of a match is unknown. This would be different if we had used a Bayesian learning approach. Then any new information that becomes available leads to an update of agents' beliefs. Also note that when the match quality is revealed, $\hat{\pi}$ takes on the value of either 1 or 0 depending on whether the match is good or bad.

determined each period by bilateral Nash bargaining between the worker and the firm¹⁴. Notice that the output realization of the current period is observed after wages have been negotiated. We do not allow for the within-period re-negotiation of wages.

With probability $(1 - \sigma)\rho_k$ the match survives to the next period. Notice that $\sigma \in [0, 1]$ is the probability that the match is hit by an exogenous separation shock. Three events can occur in-between two periods: (a) the match quality remains unknown, which happens with probability $(1 - \varphi)$ and yields payoff $J(k', i', \hat{\pi})$, (b) the match quality is fully revealed and found to be good, which happens with probability $\varphi\hat{\pi}$ and yields payoff $J(k', i', 1)$ or (c) the match quality is fully revealed and found to be bad. This happens with probability $\varphi(1 - \hat{\pi})$. Bad matches break up and thus the continuation value associated with case (c) is equal to zero. If the match survives the worker's type might change. This is taken account for by the transition probabilities given by $\mu_{i|i'}^e$.

Employed Workers

Similarly we can write the value of a job to a worker as

$$\Gamma^e(k, i, \hat{\pi}) = \max \left\{ w(k, i, \hat{\pi}) + \beta \rho_k \Gamma^+(k, i, \hat{\pi}), \sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k, i') \right\} \quad (2.6)$$

Every period the worker is free to quit his current job and to leave the firm for unemployment. A voluntary separation and the subsequent transition into unemployment might cause a change in the worker's type. Hence, the worker's outside option is given by the expected value of being unemployed $\sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k, i')$. The value of employment consists of the one period return, i.e. the current period wage income w , and the continuation value, i.e. the expected value of staying with the current employer in the next period, Γ^+ . The latter part is defined as follows

$$\begin{aligned} \Gamma^+(k, i, \hat{\pi}) = & (1 - \sigma) \sum_{i' \in \mathcal{I}} \mu_{i|i'}^e ((1 - \varphi) \Gamma^e(k', i', \hat{\pi}) + \varphi \hat{\pi} \Gamma^e(k', i', 1)) + \\ & [\varphi(1 - \sigma)(1 - \hat{\pi}) + \sigma] \sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k', i') \end{aligned}$$

- With probability $(1 - \varphi)$ nothing is learned about the match quality. The implied next period value is $\Gamma^e(k', i', \hat{\pi})$.

¹⁴The process of wage bargaining is described in detail in Appendix D.

- If the match quality is found to be good, which happens with probability $\varphi\hat{\pi}$, the value of the job changes to $\Gamma^e(k', i', 1)$. Given that the worker stays employed he might experience a type transition which is captured by $\mu_{i|i'}^e$. The same is also true for the previous case.
- If the match is revealed to be bad, which happens with probability $\varphi(1-\hat{\pi})$, or hit by an exogenous separation shock, with probability σ , the worker transits to unemployment. The value of unemployment is given by $\Gamma^u(k', i')$. Due to the break-up the worker might experience a change in his type. This is captured by the transition probabilities $\mu_{i|i'}^b$.

2.3.5 Equilibrium

All decisions, including those about match formation and break-up, are taken jointly by the firm and the worker. Hence, the criterion used in the decision making process is the joint surplus of a match. We can write it as

$$S(k, i, \hat{\pi}) = J(k, i, \hat{\pi}) + \Gamma^e(k, i, \hat{\pi}) - \sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k, i')$$

The joint surplus S consists of the sum of the value of the job to the firm and the worker, net of their respective outside values. The outside option for the worker is given by the value of unemployment and for the firm it is the value of an open vacancy (which in equilibrium is equal to zero). One can use the value functions in (2.5) and (2.6) to obtain an explicit expression for S . We state this expression in Appendix D. As mentioned previously, the firm and the worker's decision whether or not to form a match depends on the value of their common beliefs $\hat{\pi}$. We can define the reservation belief as the value of $\hat{\pi}$ for which the firm and the worker are indifferent between creating and not creating the match. We denote this value by $\underline{\hat{\pi}}$. For a given (k, i) , $\underline{\hat{\pi}}$ must fulfill

$$S(k, i, \underline{\hat{\pi}}) \begin{cases} > 0, & \forall \quad \hat{\pi} > \underline{\hat{\pi}} \\ = 0, & \text{for } \hat{\pi} = \underline{\hat{\pi}} \\ < 0, & \forall \quad \hat{\pi} < \underline{\hat{\pi}} \end{cases} \quad (2.7)$$

In what follows, we state a number of propositions that characterize some important properties of $\underline{\hat{\pi}}(k, i)$. The first establishes the existence and uniqueness of $\underline{\hat{\pi}}$ for all age cohorts k and all worker types i .

Proposition 2.1 *For each pair (k, i) , there exists a unique threshold belief $\hat{\pi}(i, k) \in (0, 1)$ if $y^g > b \geq y^b$ holds.*

The proof is provided in Appendix A. Intuitively, first we show that the surplus function $S(\cdot)$ is linear in $\hat{\pi}$, which rules out the existence of multiple $\hat{\pi}$ that are consistent with the conditions stated in (2.7). Then we establish conditions for which the joint surplus for the lowest possible value of beliefs, i.e. $\hat{\pi} = 0$, is always negative. A necessary and sufficient condition for this to hold is $b > y^b$. This is intuitive, as for $\hat{\pi} = 0$ the match produces output equal to y^b every period and thus if $b > y^b$ then the value of employment is always lower than the value of unemployment, so guaranteeing a negative match surplus. Lastly, we seek conditions that ensure $S(\cdot, \cdot, 1) > 0$. A necessary and sufficient condition is $y^g > b$. The intuition is as before. For $\hat{\pi} = 1$ the match produces output y^g with probability one. Hence, if $y^g > b$ then the value of employment always exceeds the value of unemployment, which guarantees a positive match surplus. The conditions needed for existence and uniqueness of an equilibrium also ensure that good matches persist and bad matches break-up.

Proposition 2.2 *For a given type i , the threshold belief $\hat{\pi}(i, k)$ is increasing in age k .*

For this proposition, we do not provide a fully-fledged proof but rather the intuition. Notice first that, due to the Nash bargaining, we have $J(k, i, \hat{\pi}) = (1 - e)S(k, i, \hat{\pi})$, where e denotes a firm's bargaining power. Hence, all the properties underlying $J(\cdot)$ are directly applicable to $S(\cdot)$ ¹⁵. From inspecting (2.5), we can infer that $J(\cdot)$ is essentially given by the finite sum over all $K - k$ single period returns $E(y|\hat{\pi}) - w(k, i, \hat{\pi})$, weighted by the survival probabilities, ρ_k . Each of the single period elements is necessarily positive. Therefore, the longer the horizon over which a firm can collect the profits, the larger will be the total sum $J(\cdot)$. If, however, a worker has only a few periods left in the labor market then it will be less profitable for a firm to match with that worker. $J(k, i, \hat{\pi})$ is decreasing in k , hence, if $\hat{\pi}(i, k)$ is such that $J(k, i, \hat{\pi}(i, k)) = 0$, it follows that $J(k + 1, i, \hat{\pi}(k, i)) < 0$. At the same time, we know that $\frac{\partial J(k, i, \hat{\pi})}{\partial \hat{\pi}} > 0$ (see Proof of Proposition 2.1) and so $\hat{\pi}(k + 1, i) > \hat{\pi}(k, i)$ is necessary

¹⁵Notice that the total surplus of a match is shared between the worker and the firm according to a sharing rule which is the result of a Nash bargaining process. In the bargaining process, the firm and the worker set the wage $w(i, \hat{\pi})$ in order to maximize the Nash product that is given by $[J(k, i, \hat{\pi})]^e \left[\Gamma^e(k, i, \hat{\pi}) - \sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k, i') \right]^{1-e}$. This maximization problem yields the familiar first order condition $(1 - e)J(k, i, \hat{\pi}) = e \left[\Gamma^e(k, i, \hat{\pi}) - \sum_{i' \in \mathcal{I}} \mu_{i|i'}^b \Gamma^u(k, i') \right]$. By combining the first order condition with the definition of the joint match surplus, one gets the surplus sharing rule used in the text.

to establish $J(k+1, i, \hat{\pi}(k+1, i)) = 0$. Economically speaking, this means that firms need to be compensated with a higher surplus when matching with workers that are close to the exit age. This is achieved by raising the threshold belief $\hat{\pi}$.

Proposition 2.3 *For a given age k , the reservation belief $\hat{\pi}$ is increasing in the type i*

We provide the intuition underlying this proposition a little further below, albeit for a simpler case though. This says that workers that are associated with a more precise signal, face a relatively higher threshold value than workers with a noisy signal. Next we define how the equilibrium can be characterized.

Definition 2.1 *A stationary equilibrium of the model consists of the following objects*

1. A collection of threshold beliefs $\{\hat{\pi}(k, i)\}_{k \in \mathcal{K}, i \in \mathcal{I}}$ that satisfy condition (2.7)
2. A constant mass of firm/worker matches $\Phi^+(k, i, \hat{\pi})$ for all possible combinations of $(k, i, \hat{\pi})$, and a constant mass of unemployed workers $\Phi^-(k, i)$ for all combinations of workers' type i and age k .

Before we proceed to study certain properties of the equilibrium and compute aggregate variables, we provide a brief overview of the within-period timing of the model

1. At the beginning of the period, firms and workers that were already matched in the period before observe the current state of the match $(k, i, \hat{\pi})$. They then decide whether to terminate or continue the match. Workers whose matches were dissolved enter unemployment, where they receive b .
2. Wages w are negotiated bilaterally between firms and workers.
3. Matches that continue, produce output y . Wages w get paid to workers.
4. On the basis of the current output realization, the firm and the worker update their (common) beliefs about the true quality of the match. This is done in an all-or-nothing fashion. With probability φ the true match quality is revealed and with probability $1 - \varphi$ nothing is learned. Matches that are found to be bad are dissolved.
5. Unemployed workers encounter firms with an open job with probability $\tilde{\lambda}$. Each pair that meets observes a signal γ about the quality of the match. Conditional on the signal, they form beliefs about the true probability that the match is good. Those beliefs are expressed by $\hat{\pi} \equiv E(\pi|\gamma, i)$. On the basis of $\hat{\pi}$ they decide whether or not to form a match. New matches start producing output in the next period.

6. At the end of the period, workers' age and experience change in their type i depending on their labor market state.
7. Workers die with probability $(1 - \rho_k)$, jobs get hit by the destruction shock with probability σ and new workers, aged $k = 0$, enter the market.

2.4 *A Simplified Version*

The model we have laid out in the previous section includes, subjectively, the most important features needed to capture the inter-temporal aspect of an individual's working life cycle. In particular the structure we impose allows us to describe how an individual's labor market performance during early periods in the working life cycle spills over to the job finding and hiring conditions they face at a later stage in life. The inter-temporal aspect is captured by the the notion of worker-type-specific information imperfections in the hiring process that change as the individual builds up a particular labor market history. The model can, therefore, account for the transition process by which individuals move through various stages of their working life cycle, where each of those stages is associated with different conditions regarding job finding, job turnover, unemployment and the like.

Before we proceed to the empirical part of the paper, it would be of interest to study how the information imperfections, captured by the type- i -dependent noise component in the initial signal, affect the labor market conditions of workers of different types. By labor market conditions, we mean variables on the individual level such as job finding and turnover rates but also aggregate outcomes, such as type-specific unemployment rates. The modeling framework as it is now, however, is too complex to address these issues by analytical means. The complexity stems mainly from the life-cycle component which is captured by the type transition and the age dependent survival probabilities.

In this section of the paper, we consider a simplified version of the original model. The simplification allows us obtain closed form solutions for individual and aggregate outcomes and so we can explicitly study the role of worker-specific information imperfections in the hiring process and how they shape aggregate outcomes, such as job turnover and unemployment. The simplification is achieved by abstracting from the life cycle component, i.e. we set (a) $\rho_k = \rho \ \forall k \in \mathcal{K}$, (b) $\mu_{i|i'} = 0 \ \forall i' \neq i$ and (c) $\mu_{i|i'} = 1$ for $i' = i$. By the first condition, we eliminate the time dependent survival probabilities and instead, we introduce a constant probability of death. By the latter two, we abstract from the process that governs changes

in workers' i -type. Instead, we assume that a worker's type remains the same throughout his lifetime.

2.4.1 Equilibrium

By implementing these simplifications, we arrive at the following expressions for the value of unemployment and the joint surplus of a match.

$$\Gamma^u(i) = b + \beta\rho\tilde{\lambda} \int_{\hat{\pi}} \max\{\Gamma^e(i, \hat{\pi}) - \Gamma^u(i), 0\} dG^i(\hat{\pi}) + \beta\rho\Gamma^u(i) \quad (2.8)$$

$$S(\hat{\pi}, i) = \max\{E(y|\hat{\pi}) + \beta\rho(1 - \sigma)[\varphi\hat{\pi}S^g(i) + (1 - \varphi)S(\hat{\pi}, i)] - (1 - \beta\rho)\Gamma^u(i), 0\} \quad (2.9)$$

The first step in analyzing the equilibrium of the simplified model is to find the threshold beliefs $\hat{\pi}(i)$. To this end we make use of the linearity of $S(\hat{\pi}, i)$ in $\hat{\pi}$ and the condition $S(i, 1) = S^g(i)$. This allows us to express the surplus function as $S(\hat{\pi}, i) = (\hat{\pi} - \underline{\hat{\pi}})S'(\underline{\hat{\pi}})$, where

$$S'(\underline{\hat{\pi}}, i) = \frac{y^g - y^b}{1 - \beta\rho(1 - \sigma)(1 - \varphi\underline{\hat{\pi}})}$$

Setting $S(\hat{\pi}, i) = 0$ in Equation (2.9) and using the surplus sharing rule $\Gamma^e(i, \hat{\pi}) - \Gamma^u(i) = (1 - e)S(i, \hat{\pi})$ together with Equation (2.8) we can write the equilibrium condition as follows

$$\underline{\hat{\pi}}S'(\underline{\hat{\pi}})[1 - \beta\rho(1 - \sigma)(1 - \varphi)] + y^b = b + \beta\rho\tilde{\lambda}(1 - e)S'(\underline{\hat{\pi}}) \int_{\hat{\pi} \geq \underline{\hat{\pi}}} (\hat{\pi} - \underline{\hat{\pi}}) dG^i(\hat{\pi}) \quad (2.10)$$

Obviously, the $\underline{\hat{\pi}}$ that solves Equation (2.10) is a candidate for an equilibrium.

Proposition 2.4 *An interior equilibrium exists and is unique iff the following conditions are met*

- $y^g > b$
- $b + \beta\rho\tilde{\lambda}(1 - e)\bar{\alpha}S'(0,) > y^b$

The proof is stated in Appendix A. Having found the equilibrium threshold beliefs, we can ask the question of how those beliefs differ across i -types. This question is important insofar as the threshold value for a particular i -type determines the hiring conditions of type i workers. Using a simple graphical analysis, we find that the right-hand side of Equation (2.10) can be represented by a downward sloping curve which, for the same value of $\hat{\pi}$, shifts to the right for $i' > i$. At the same time, the upward-sloping curve that represents the left-hand side stays the same for all worker types, as it is a function only of $\hat{\pi}$. From that it follows that $\hat{\pi}(i') > \hat{\pi}(i)$ for $i' > i$, implying that the equilibrium cutoff value is lower for low type workers. Recall that "low type workers" are those that emit a more noisy signal in the hiring process.

Notice first that: the right hand side of Equation (2.10) is the only place where the type enters the equilibrium condition and, moreover, it is directly related to the expected surplus of a match for an unemployed worker. Low type workers have a lower expected surplus. This is due to the fact that the equilibrium distribution of their beliefs is less dispersed than that of high type workers¹⁶. Given that the range of acceptable draws is truncated below by $\hat{\pi}$, this results in a lower expected value for low type workers. From Equation (2.8) we know that this reduces the value of unemployment for a worker and thereby makes the outside option of an employed worker less valuable. A lower outside option implies that a match is profitable for lower draws of $\hat{\pi}$, which is due to the fact that those workers require less compensation for making them leave unemployment.

2.4.2 Threshold signals and threshold beliefs

In the previous section, we have established the result that high noise workers face a lower threshold belief than workers with a precise signal. But what can be said about the relation between the actual signals γ across worker types that induce the respective threshold beliefs? Recall that the same signal γ leads to a different $\hat{\pi}$ depending on the worker's type i . Therefore, we can ask the question how the minimum signal $\gamma \rightarrow \hat{\pi}$ compares across types? For brevity we move the discussion of this issue to Appendix B. Here we continue with the analysis of aggregate outcomes

¹⁶Notice that this is a direct consequence of the way how beliefs are formed. From Condition 2.2 we know that for high (low) type workers more (less) weight is put on the actual signal and less (more) weight is put on the mean value α . Therefore, the resulting beliefs are more dispersed (more centered around the mean).

2.4.3 Aggregate Outcomes

Type-specific unemployment rates

In this section we focus on aggregate equilibrium outcomes and in particular, we consider three different objects: (1) the type-specific unemployment rate, (2) a measure of job turnover and (3) a measure of individuals' job stability. For expositional convenience we state here only the most important results. The calculations that lead to these results can be found in Appendix C. Let u^i denote the mass of unemployed, type i workers. The share of each type in the population is given by I^i , hence $\sum_{i=1}^I I^i = 1$. We can write the type-specific unemployment rate as

$$\tilde{u}^i = \frac{u^i}{I^i} = \frac{1}{1 + \tilde{a}[1 - G^i(\hat{\pi}^i) + \tilde{c} \int_{\hat{\pi}^i} \hat{\pi} g^i(\hat{\pi}) d\hat{\pi}]} \quad (2.11)$$

where $\tilde{a} = \frac{\tilde{\lambda}\rho}{1-(1-\sigma)\rho(1-\varphi)}$ and $\tilde{c} = \frac{\varphi(1-\sigma)\rho}{1-\rho(1-\sigma)}$. What can we say about type-specific unemployment rates? Which workers, those with noisy signals or those with precise information, are more affected by unemployment? To address this issue we can exploit the dependence of the hiring threshold $\hat{\pi}$ on the worker's type i . This allows us to write the unemployment rate as $\tilde{u}^i = \tilde{u}^i(\hat{\pi}^i(\sigma_{\epsilon,i}^2), \sigma_{\epsilon,i}^2)$. Computing the change in unemployment in response to more noise in the signal, $\frac{d\tilde{u}^i}{d\sigma_{\epsilon,i}^2}$, therefore becomes a simple application of the chain rule, i.e. $\frac{d\tilde{u}^i}{d\sigma_{\epsilon,i}^2} = \frac{\partial \tilde{u}^i}{\partial \hat{\pi}^i} \frac{\partial \hat{\pi}^i}{\partial \sigma_{\epsilon,i}^2} + \frac{\partial \tilde{u}^i}{\partial \sigma_{\epsilon,i}^2}$. Increasing the amount of noise exhibits a direct effect on unemployment, represented by the second term, and also an indirect effect through the change in the hiring threshold, $\hat{\pi}$. Computing the differential change gives

$$\frac{d\tilde{u}^i}{d\sigma_{\epsilon,i}^2} = \tilde{a}\tilde{u}^i \left[(1 + \tilde{c}\hat{\pi}^i) g(\hat{\pi}^i) \frac{\partial \hat{\pi}^i}{\partial \sigma_{\epsilon,i}^2} - \left(-\frac{\partial G^i(\hat{\pi}^i)}{\partial \sigma_{\epsilon,i}^2} + \tilde{c} \int_{\hat{\pi}^i} \hat{\pi} \frac{\partial g^i(\hat{\pi})}{\partial \sigma_{\epsilon,i}^2} d\hat{\pi} \right) \right] \quad (2.12)$$

Both terms inside the square brackets have an intuitive economic interpretation. The first term relates to the change in the flow of workers out of unemployment. As shown previously, more noise in the signal leads to a decrease in the critical threshold, $\hat{\pi}$, represented by $\frac{\partial \hat{\pi}^i}{\partial \sigma_{\epsilon,i}^2} < 0$. This facilitates the matching of workers to firms since the minimum signal that is required for a successful match creation falls. A drop in the threshold signal will make workers exit unemployment more quickly and thereby, it reduces the duration and also the level of unemployment. The size of the reduction in unemployment is proportional to the mass of workers that are located at the threshold and are, therefore, affected by the differential change in the threshold. This mass is related to the first term inside the brackets given by $(1 + \tilde{c}\hat{\pi}^i) g(\hat{\pi}^i)$.

The second term in Equation (2.12) measures the direct effect of more noise on unemployment. This term consists of two parts. The first relates again to the change in the flow of workers out of unemployment. However, this time the hiring threshold $\hat{\pi}$ is kept constant which allows us to capture the variation in the flow that can be attributed solely to a change in the noisiness of workers' signals. This term, $-\partial G^i(\hat{\pi}^i)/\partial \sigma_{\epsilon,i}^2$, is negative whenever $\hat{\pi}(i) > \alpha$ and is positive when the inequality is reversed. Let's first focus on the case when $\hat{\pi}(i) > \alpha$. More noise in the signal changes the way agents form their beliefs about the true probability that the match is good π . In particular, when the signals become less informative agents rely to a lesser extent on the actual signal and instead they put more weight on the mean of π . Therefore, the same signal leads to lower beliefs $\hat{\pi}$. Consequently, it becomes harder for agents to generate a signal that exceeds the threshold. Any meeting between a firm and a worker is, therefore, less likely to result in a successful match creation. This impedes the flow of workers out of unemployment and increases unemployment.

The second term inside the square brackets relates to the change in the amount of mismatch. As mentioned previously, more noise in the signal leads to beliefs that are more located around the mean. Consequently, the distribution becomes less dispersed and the expected value of beliefs above the threshold $E(\hat{\pi} > \hat{\pi})$ drops. $E(\hat{\pi} > \hat{\pi})$ is a natural measure of the average quality of existing matches. A drop of which indicates that there are relatively fewer matches in place that were formed on the basis of good beliefs. Consequently, the amount of mismatch among existing matches rises. Mismatch leads to a separation, hence a larger fraction of bad matches leads to a rise in the flow of workers into unemployment. The associated rise in unemployment is represented by the term $\int_{\hat{\pi}} \hat{\pi} \frac{\partial g^i(\hat{\pi})}{\partial \sigma_{\epsilon,i}^2} d\hat{\pi}$ which is negative for $\hat{\pi} > \alpha$.

The total effect of more noise on unemployment is ambiguous as it depends on the location of the threshold $\hat{\pi}$. The first effect described by $\frac{\partial \tilde{u}^i}{\partial \hat{\pi}^i} \frac{\partial \hat{\pi}^i}{\partial \sigma_{\epsilon,i}^2}$ is always negative. However, the second effect $\frac{\partial \tilde{u}^i}{\partial \sigma_{\epsilon,i}^2}$ is positive for $\hat{\pi}(i) > \alpha$ but negative when the inequality is reversed.

Match destruction

Another object of interest is the match separation rate. It is computed as the ratio of bad matches to the total number of matches with unknown quality. Due to learning, bad matches eventually dissolve. Thus the fraction of bad matches will be indicative for the prevailing

degree of job turnover in the economy. More importantly, when comparing the measure for different i types, we can draw conclusions about which type of worker is more affected by job turnover and thus, enjoys less stable jobs. Let Φ^i denote the match destruction rate that prevails in equilibrium. We can write as

$$\Phi^i = 1 - E^i(\hat{\pi}|\hat{\pi} \geq \hat{\pi}) \quad (2.13)$$

$E^i(\hat{\pi}|\hat{\pi} \geq \hat{\pi})$ denotes the (conditional) mean of beliefs across existing matches and is an insightful measure for the average quality of the signals on which matches are based upon. It is particularly insightful because on the basis of it we can assess (a) the degree of mismatch across existing firm/worker pairs and (b) how mismatch compares across workers of different i -types. We find that $dE^i(\hat{\pi}|\hat{\pi} \geq \hat{\pi})/d\sigma_{\epsilon,i}^2 < 0$ from which it follows that $d\Phi^i/d\sigma_{\epsilon,i}^2 > 0$. It says that workers with a less informative signal are, on average, affected more by job turnover than workers with a precise signal. This is intuitive as when there is more noise in the signal, then agents are less likely to generate high beliefs. Hence, the expected value of $\hat{\pi}$, conditional on the signal being greater than the threshold, drops. Further, the average quality of beliefs is lower as fewer matches are based on good beliefs and more mass is located around the threshold. This is the same as saying that the amount of mismatch among existing matches rises. Mismatch, once detected, leads to a separation/break-up, thus workers with less informative signals are more affected by job turnover as their jobs are, on average, less stable.

Job stability

Regarding job stability, we can compute the ex-ante probability of a worker experiencing a job destruction. This can be written as

$$P(destr) = \sigma + (1 - \sigma)\varphi[1 - E^i(\hat{\pi}|\hat{\pi} \geq \hat{\pi})] \quad (2.14)$$

There are two sources of job destruction. A match can be hit by an exogenous destruction shock, which happens at the Poisson rate σ , or it can be dissolved due to mismatch, this is captured by the second part. The probability of experiencing an endogenous separation is given by the ex-ante probability of ending up in a bad match. The expected duration of a match which is given by $P(destr)^{-1}$. We saw previously that $dE^i(\hat{\pi}|\hat{\pi} \geq \hat{\pi})/d\sigma_{\epsilon,i}^2 < 0$. Thus, we can establish the result that agents with less reliable signals face a higher probability of experiencing a job termination.

2.5 Solving the Full Model

This section outlines the solution algorithm that was used to compute the equilibrium of the full model presented in Section 2.3. The state of an unemployed workee is given by (k, i) and that of a match is $(k, i, \hat{\pi})$. Notice that k (age) and i (type) are both discrete objects. The aging process in the model is constructed in a way so that an individual with age k today is of age $k' = k + 1$ next period. Furthermore, the process that governs the type transition is a discrete valued Markov process implying that the realizations of i all lie on a pre-specified grid. This suggests the use of equally-spaced grids for k and i which we specify as follows: $\mathcal{K} = \{1, 2, \dots, \bar{K}\}$ and $\mathcal{I} = \{1, 2, \dots, \bar{I}\}$. \bar{K} is the maximum number of periods an individuals stays in the labor market and \bar{I} is the highest achievable type. For the beliefs $\hat{\pi}$, we adopt a continuous representation of the state space.

The value function of an unemployed worker and the joint surplus function of a match are described by the mappings $\Gamma : \mathcal{K} \times \mathcal{I} \rightarrow \mathbb{R}$ and $S : \mathcal{K} \times \mathcal{I} \times [0, 1] \rightarrow \mathbb{R}$. Due to the discreteness of its domain, we can express Γ as a collection of points $\{\Gamma_{k,i}\}_{k \in \mathcal{K}, i \in \mathcal{I}}$. S , on the other hand, needs to be approximated. We do not use the same method to approximate S on the entire $[0, 1]$ domain but, instead, we combine two different methods. The reason for that is the kink in the value function that is induced by the cutoff value $\hat{\pi}$. In particular we have $S(k, i, \hat{\pi}) = 0 \forall \hat{\pi} \leq \hat{\pi}$ and $S(k, i, \hat{\pi}) > 0 \forall \hat{\pi} > \hat{\pi}$. To accommodate for that we use a piecewise-linear approximation of S for values of $\hat{\pi} \in [0, \hat{\pi})$ and a Chebyshev polynomial for $\hat{\pi} \in [\hat{\pi}, 1]$. For each node on the $\mathcal{K} \times \mathcal{I}$ grid, we construct the polynomial by running a Chebyshev regression using 50 collocation nodes. These nodes are given by the collection $\{z_{k,i,x}\}_{k \in \mathcal{K}, i \in \mathcal{I}, x \in \{1, 2, \dots, 50\}}$.

The value functions Γ and S are found jointly using an algorithm that incorporates a fixed-point iterative scheme. The algorithm is structured in the following manner:

1. The first step involves guessing a joint surplus function S . The initial guess is denoted by S^0 .
2. The second step involves computing the threshold beliefs. For each combination of (k, i) we identify the $\hat{\pi}$ for which $S(k, i, \hat{\pi}) = 0$.
3. For each pair (k, i) we allocate the collocation nodes in the interval $[\hat{\pi}(k, i), 1]$ and run the Chebyshev regression to find the approximation coefficients associated with S . Those coefficients are given by $\{c_{k,i,x}\}_{k \in \mathcal{K}, i \in \mathcal{I}, x \in \{1, 2, \dots, 50\}}$.

4. Next, we solve for $\{\Gamma_{k,i}\}_{k \in \mathcal{K}, i \in \mathcal{I}}$ by first using the terminal condition $\Gamma(\bar{K}, i) = b$ for $\forall i$ and then working backwards. The integral in Equation (2.3) is computed using a Gauss-Legendre quadrature with 7 quadrature nodes.
5. To compute $\{S_{k,i}^g\}_{k \in \mathcal{K}, i \in \mathcal{I}}$ it suffices to evaluate $S(k, i, \hat{\pi})$ at $\hat{\pi} = 1, \forall k, i$. This follows from $S^g(k, i) = S(k, i, 1)$.
6. To solve for the new S^{j+1} we proceed as follows: We use (a) the value of the previous iteration step S^j , (b) the terminal condition $S(\bar{K}, i, z) = 0, \forall i, z$, (c) the value of unemployment $\{\Gamma_{k,i}\}_{k \in \mathcal{K}, i \in \mathcal{I}}$ and work backwards for all (k, i, z) .
7. Lastly, we evaluate $\|S^{j+1} - S^j\|$. If the distance is smaller than some convergence criterion we terminate the algorithm, else we set $S^j = S^{j+1}$ and return to step 2.

2.6 The Estimation Strategy

The structural model is estimated by indirect inference. This method, which was first proposed by Gourieroux, Montfort and Renault (1993), belongs to the class of simulation-based estimation procedures¹⁷. Estimators of this sort are particularly convenient in situations when the complexity of the model leads to an intractable likelihood function or when some of the variables are unobserved in the data. The idea of indirect inference as it is used here is to match the statistical properties of the simulated data with those of observed data along selected dimensions. The dimensions, along which the simulated and the observed data are evaluated, are represented by an auxiliary model. The central idea is then to choose the parameters of the structural model in a way so that the parameters of the auxiliary model estimated from actual and from simulated data are as close as possible.

Obviously, a crucial step in the estimation procedure concerns the selection of the auxiliary model. A "good" auxiliary model is one that captures well the statistical properties of the data along the dimension(s) required to identify the model's structural parameters. The parameters estimated from the auxiliary model have to contain enough information, in the sense that they have to be sufficiently responsive to changes in the model's structural parameters. If that was not the case then the objective function would exhibit relatively little

¹⁷Similar estimation methods include efficient method of moments (Gallant and Tauchen (1996)), or simulated (quasi-)maximum likelihood (Smith (1993)), For a discussion of simulation-based estimators see Tauchen (1997) and Gourieroux and Monfort (1996).

curvature and the structural parameters would be poorly identified. In this respect, the suitability of an auxiliary model can be judged by looking at the standard errors of the estimated of the structural parameters. Secondly, the auxiliary model should be fast to compute. The estimation of the structural parameters typically requires a high number of evaluations of the model's objective function. This can potentially be very time consuming. Hence, choosing a parsimonious auxiliary model is key for keeping the computational time at a reasonable level.

As the auxiliary model, I choose a discrete-time hazard model that is tailored to the analysis of duration data. The choice is motivated by the nature of the economic process that we are looking at. The structural model is designed to capture differences in job duration across age cohorts. Therefore, a hazard rate model is a natural candidate since it allows for identification of the effect of individuals' characteristics on the actual job duration.

The data that is used in the estimation consists of observations of individuals' employment spells. Each observation is indexed by $j \in \{1, 2, \dots, J\}$, where J denotes the total number of observations. For the estimation I divide the time line into N intervals that are given by $[0, \tau_1)$, $[\tau_1, \tau_2)$, ..., $[\tau_{N-1}, \tau_N)$. Intervals are indexed by the subscript $n \in \{1, 2, \dots, N\}$. Let t_j denote the duration of employment spell j and let $y_{j,n}$ be a binary variable that is zero if in interval n the worker is still employed in job j and unity otherwise. Thus, $y_{j,n}$ will be zero for all intervals n for which $\tau_n < t_j$ and it will be one for all intervals or which $\tau_n \geq t_j$. For each employment spell we observe a string of zeros followed by a string of ones¹⁸. The important information is the time interval for which $y_{j,n}$ becomes unity for the first time. This is because the switch from zero to one indicates a match separation. Notice, however, that some observations are right-censored. That is, we can not determine the exact time when the job ended. This can be due to two reasons: (a) the duration of the employment spell t_j exceeds the time that is captured by the intervals, which is the case whenever the job lasted longer than the period that is captured by the N intervals, i.e. $t_j \geq \tau_N$, or (b) if the employment spell is still ongoing at the time the latest survey was taken (2006). In the first case, all observations are right-censored at τ_N whereas in the second case, the observation is right-censored in the interval corresponding to the job duration at the time of the survey. To account for censoring, we construct the binary variable c_j that is unity when the duration of job j is uncensored and zero otherwise. Notice that for the estimation it will be important whether the ending of an employment spell represents a true job loss or is due to censoring. This information will be captured by c_j .

¹⁸Notice that $y_{j,n} = 1$ implies $y_{j,n+1} = 1$.

The hazard rate of separating from job j in interval n is modeled as a piecewise constant function given by

$$\lambda(t; \mathbf{x}_j, \beta) = \kappa(\mathbf{x}_j, \eta) \lambda_n, \quad t \in [\tau_{n-1}, \tau_n) \quad (2.15)$$

β is a vector of unknown parameters and $\kappa(\mathbf{x}_j, \eta) > 0$ is a function of observable worker characteristics. Those characteristics are captured by the vector \mathbf{x}_j , whereas η is a vector of unknown parameters. The baseline hazard for each interval is given by λ_n . This specification implies a constant hazard within a particular interval n , but it allows the hazard to differ across intervals. Notice that the existence of a baseline hazard makes a constant term as a covariate superfluous. For a particular employment spell j only the intervals for which $y_{j,n} = 0$ can be used to predict the probability of $y_{j,n+1}$ being zero or one in the subsequent interval $n + 1$ ¹⁹. These (conditional) probabilities are given by

$$\begin{aligned} P(y_{j,n} = 1 | y_{j,n-1} = 0, \mathbf{x}_j) &= P(\tau_{n-1} \leq t_j < \tau_n | t_j \geq \tau_{n-1}, \mathbf{x}_j) \\ &= 1 - \exp \left[- \int_{\tau_{n-1}}^{\tau_n} \lambda(s; \mathbf{x}, \beta) ds \right] \\ &= 1 - \exp [-\kappa(\mathbf{x}_j, \eta) \lambda_n (\tau_n - \tau_{n-1})] \end{aligned} \quad (2.16)$$

$$P(y_{j,n} = 0 | y_{j,n-1} = 0, \mathbf{x}_j) = \exp [-\kappa(\mathbf{x}_j, \eta) \lambda_n (\tau_n - \tau_{n-1})]$$

Using those probabilities one can construct the log likelihood function. If for employment spell j the separation - censored or uncensored - occurred in interval n_j the log likelihood is given by

$$\begin{aligned} l(n_j, \mathbf{x}_j, c_j; \beta) &= - \sum_{h=1}^{n_j-1} \kappa(\mathbf{x}_j, \eta) \lambda_h (\tau_h - \tau_{h-1}) \\ &\quad + c_j \log \{ 1 - \exp [-\kappa(\mathbf{x}_j, \eta) \lambda_{n_j} (\tau_{n_j} - \tau_{n_j-1})] \} \end{aligned} \quad (2.17)$$

The first part is the probability the job lasts until n_j and the second part is the probability that it ends in the interval n_j . Notice that the latter part is non-zero only when the observation is uncensored and therefore represents a true job loss. The log likelihood function for the entire sample is obtained by summing over $j = 1, \dots, J$ employment spells. It is given by

$$\mathcal{L}(\mathbf{y}_J; \beta) = \sum_{j=1}^J l(n_j, \mathbf{x}_j, c_j; \beta) \quad \text{where} \quad \mathbf{y}_J = \{n_j, \mathbf{x}_j, c_j\}_{j=1}^J \quad (2.18)$$

¹⁹This is due to the fact that once $y_{j,n} = 1$ we know that $y_{j,n+1} = 1$.

At this point let us discuss the worker characteristics that are considered and the functional form of $\kappa(\cdot, \cdot)$. For a particular employment spell j we use information on: (a) a_j : the worker's age at the time when the match was formed and (b) ω_j : the percentage deviation of the worker's initial wage in job j from the average initial wage paid in matches that were formed by workers of the same age. According to the structural model a_j and ω_j both contain important information about an individual's hazard rate. Recall that a worker's age at match formation is a good predictor of his labor market experience which in turn is indicative for the worker's type. The type determines the precision of the signal that is emitted in the hiring process and it thereby affects the survival probability of the employment relationship. A worker's age at match formation will, therefore, be informative about the likelihood of separation.

The channel through which ω_j affects the hazard rate goes via agents' beliefs. Agents are rational in the sense they have correct beliefs about the probability that the match is good²⁰. Consequently, agents' beliefs are inversely related to the **actual** likelihood of match separation. The probability of survival determines the joint surplus of a match. More durable matches generate a higher surplus to the worker and the firm than matches which are expected to break-up soon. Typically the value of the job is reflected by the worker's remuneration, i.e. the wage. As a result, a positive deviation of the worker's initial wage from the average initial wage, reflected by a positive ω_j , is an indication that agents expect their employment relation to be stable and to yield a high return.

We do not treat the workers age at match formation as a continuous variable. Instead, we consider age intervals and work with dummy variables. In particular, we divide the "age-line" into $K + 1$ distinct intervals that are given by $[a_0, a_1)$, $[a_1, a_2)$, ..., $[a_{K-1}, a_K)$, $[a_K, \infty)$. Let \tilde{a}_k be a dummy variable that is unity when the worker's age at match formation falls in the interval $[a_{k-1}, a_k)$, and zero otherwise. The vector of covariates \mathbf{x}_j is thus given by $(\tilde{a}_1, \dots, \tilde{a}_K, \omega_j)$. As a functional form for $\kappa(\cdot, \cdot)$ we choose

$$\kappa(\mathbf{x}_j, \eta) = \exp(\eta_{a_0|a_1}\tilde{a}_1 + \eta_{a_1|a_2}\tilde{a}_2 + \dots + \eta_{a_{K-1}|a_K}\tilde{a}_K + \eta_\omega\omega_j)$$

The age category that serves as a means of comparison is given by the interval $K + 1$, i.e. $[a_K, \infty)$. Hence, any given age coefficient $\eta_{a_{k-1}|a_k}$ $k \in \{1, 2, \dots, K\}$ has to be interpreted as the difference in the hazard between a worker that starts a new job at age $a \in [a_{k-1}, a_k)$ and

²⁰The term "correct" should reflect that agents' beliefs are, on average, consistent with the true match quality.

a worker that starts a job at age $a \geq a_K$, everything else equal. If those coefficients were all the same across age cohorts then, in terms of separation probability, it would not matter at what age a worker starts a new job.

Using the information on (n_j, \mathbf{x}_j, c_j) that is available for each employment spell j , we can estimate the coefficients of the auxiliary model by maximizing the log likelihood function in (3.3). The vector of coefficients we estimate are given by the vector

$$\beta = (\eta_{a_0|a_1}, \dots, \eta_{a_{K-1}|a_K}, \eta_\omega, \lambda_1, \dots, \lambda_N)$$

The maximum likelihood estimate of β is

$$\hat{\beta}_J = \arg \max_{\beta} \mathcal{L}(\mathbf{y}_J; \beta) \quad (2.19)$$

Recall that the ultimate goal of the estimation process is to determine the parameters of the structural model. Estimating the coefficients of the auxiliary model from data can be considered as an intermediate step that is a convenient way to condense the information that is contained in the data into something very tractable, i.e. the coefficients of the auxiliary model.

We now turn to the next step in the estimation procedure. Let $\mathbf{y}_S(\phi)$ denote a matrix of S simulated observations of the endogenous variables $\{n_s, \mathbf{x}_s, c_s\}_{s=1}^S$ using the structural model and a set of structural parameters given by ϕ . Let $\beta(\mathbf{y}_S, \phi)$ denote the coefficients of the auxiliary model estimated from simulated data $\mathbf{y}_S(\phi)$. Estimating the parameters of the structural model essentially amounts to finding the value of ϕ so that the distance between the coefficients of the auxiliary model estimated from simulated data and from actual data is minimized. Technically, ϕ is chosen so that

$$\tilde{\phi}_S^H(\mathcal{Q}) = \arg \min_{\phi} \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^H \beta^h(\mathbf{y}_S, \phi) \right) \times \mathcal{Q} \times \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^H \beta^h(\mathbf{y}_S, \phi) \right)' \quad (2.20)$$

where \mathcal{Q} is a symmetric, non-negative definite weighting matrix. Gouriéroux, Montfort and Renault (1993, Proposition 4) shows that the optimal weighting matrix in this case is given by $\mathcal{Q}^* = \mathcal{Z}_0^{-1}$ where

$$\mathcal{Z}_0 = \lim_{J \rightarrow \infty} V \left[\sqrt{J} \frac{\partial \mathcal{L}(\mathbf{y}_J, \beta_0)}{\partial \beta} \right]$$

Gourieroux, et al. (1993) shows that under usual regularity conditions the estimator $\tilde{\phi}_S^H(\mathcal{Q})$ is asymptotically normal, when H is fixed and S goes to infinity:

$$\sqrt{S} \left(\tilde{\phi}_S^H(\mathcal{Q}^*) - \phi_0 \right) \rightarrow^d N[0, \mathcal{W}(H, \mathcal{Q}^*)]$$

where the variance-covariance matrix $\mathcal{W}(H, \mathcal{Q}^*)$ is given by

$$\mathcal{W}(H, \mathcal{Q}^*) = \left(1 + \frac{1}{H} \right) [\mathcal{Z}_1' \mathcal{Q}^* \mathcal{Z}_1]^{-1} \quad (2.21)$$

The matrices \mathcal{Z}_0 and \mathcal{Z}_1 are impossible to compute. However, Gourieroux, et al. (1993) shows that both can be consistently estimated by evaluating the Hessian using, respectively, observed and simulated data.

$$\mathcal{Z}_0 = p \lim_J J \frac{\partial^2 \mathcal{L}(\mathbf{y}_J; \hat{\beta}_J)}{\partial \beta' \partial \beta} \quad \mathcal{Z}_1 = p \lim_S \frac{\partial^2 \mathcal{L}(\mathbf{y}_S(\tilde{\phi}_S^H(\mathcal{Q}^*)); \hat{\beta}_J)}{\partial \beta' \partial \phi} \quad (2.22)$$

2.7 The Data

The data I use in the estimation comes from the National Longitudinal Survey of Youth 1979 (*NLSY* 79). The *NLSY* 79 is a representative sample of 12,686 young men and women who were 14 to 22 years of age when first surveyed in 1979. The cohort was interviewed annually through 1994. Since 1994, the survey has been administered biennially. The last survey that is available was published in 2006. The entire data set I use in the estimation thus consists of 27 individual surveys. The *NLSY* 79 contains demographic variables, labor market data and information on individuals' wealth and consumption.

The advantageous feature of the *NLSY* 79 is it's panel dimension. Each individual in the survey is observed from the time when they first enter the labor market (conditional on not dropping out). This allows me to construct the entire labor market history of an individual using the information on the individual's labor market transitions. It might seem problematic though that the survey is taken annually (and from 1994 even biennially) considering that a substantial fraction of jobs, especially for young individuals, last for less than a year. However, the survey is constructed in a way so that during each interview the individual is asked not just about any ongoing job but also about up to 5 employment spells that occurred since the last interview. Therefore, it is possible to recover the complete labor market history of each individual in the survey.

The data set I construct consists of employment spells of individuals, each denoted by j , and a set of individuals characteristics associated with the respective employment spell. As mentioned previously, each observation j consists of information about (a) t_j : the worker's total tenure on the job, (b) c_j : an indicator that is unity when the duration of job j is uncensored and zero otherwise, (c) a_j : the worker's age when the match was formed and (d) ω_j : the percentage deviation of the worker's initial wage in job j from the average initial wage paid in matches that were formed by workers of the same age. To guarantee a reasonable degree of homogeneity across individuals, I exclude certain observations from the data set. In particular, the final data set consists of observations of white males that had their first job at the age of 18 or later. Moreover, I exclude part time jobs and consider just those employment spells for which individuals worked for ≥ 30 hours per week²¹.

In constructing the data, particularly in determining the length of a spell of employment, I take a retrospective approach. More precisely, every time I observe a job separation in the data I use the information on the worker's total tenure on that job to determine the length of employment. Individuals are always asked to report the total number of weeks they have been working for a particular employer, irrespective of whether the job has already ended or not. Therefore, at the time of the separation I read off the worker's tenure on the job to determine the duration of the job²². For the jobs that are still ongoing at the time of the last survey (2006) we do not observe a separation. Therefore, I record the duration of the job at the time of the last survey and mark the observation as right censored.

The next step concerns the construction of a variable that, for each spell of employment j , indicates the worker's age at the time of the match formation. To that end, I use information on the worker's age at the time of the job separation and the worker's total tenure on the job. However, using simply the worker's age measured in years which is reported during

²¹The decision to exclude part-time jobs is based on the following consideration. It seems likely that firms' hiring practices and workers' job selection decisions are based on very different criteria depending on whether a part-time or a full-time job is concerned. Full-time jobs naturally involve a higher degree of commitment from both sides and hence, the decision to form a match might differ between both types of jobs.

²²The two main advantages of this approach are that (a) I can determine the exact length of each employment spell and (b) I avoid the possible double-counting of employment spells. An alternative approach would be to go the other way around, i.e. to follow a job from the beginning to its end and to record the tenure at each consecutive survey for which the job still exists. This, however, would be prone to error because any temporary interruption of an individual's record due to the non-availability for the interview, for instance, would mistakenly be considered as the ending of an employment spell. The double counting of a single spell would be unavoidable.

the interview is not advisable. This is mainly because, for our purposes a "year" as a unit of measurement is too coarse to capture the worker's exact age at the time of the match formation. To minimize the error I, thus, compute the worker's exact age (in months) at the time of the interview. This is done by using information on the month and the year of the current interview and the month and the year of the individual's birth. By subtracting the worker's tenure on a particular job (measured in months) from his current age, one gets the individual's precise age when he started the job.

The last step concerns the construction of a variable which measures the percentage deviation of the worker's initial wage in job j from the average initial wage paid in all matches that were formed by workers of the same age. The construction of this variable is a little tricky mainly because the initial wage needs to be recovered for each employment spell individually. To this end I start, for each employment spell, at the time of separation. In step (1), I use information on (a) the week number of the current interview, (b) the week number of the last interview and (c) the total tenure of worker on the current job. By using (a) and (b) we can determine how many weeks have passed since the last interview. In step (2), I consult the preceding survey and check whether the very same job has a duration of less or equal than 52 weeks. If this is the case, I record the reported wage as the worker's initial wage in that job. If the duration is more than 52, I return to step (1) and continue until the initial wage is found.

This method has two main advantages: first it can easily handle the switch from annual to biennale interviews without losing information. This is because it uses the number of weeks between two consecutive interviews to identify the survey in which one has to look for the respective job. Second, for the very same reason, it can manage a temporary discontinuation of the observability of a job which might be due to the non-availability of the interviewee.

Moreover, it is fairly precise. For 79.14% of the employment spells I am able to recover the initial wage. The failure to get a complete assignment of initial wages to jobs is due to a variety of reasons, the most important of which is the existence of missing values for wages. By inspecting the original data it becomes evident that individuals frequently do not report their wages, especially for jobs with a short duration. Arguably, excluding those observations with missing values from the data set would introduce a bias. To solve this problem, I choose to use imputed values to find the remaining initial wages. The remaining steps in constructing ω_j are straightforward. For each age cohort I compute the average initial wage and ω_j

is then simply the percentage deviation of the initial wage of each observation from the mean.

Number of employment spells, by duration (in years) and by age at match formation								
Age / Duration	[0-0.5)	[0.5-1)	[1-2)	[2-3.5)	[3.5-5)	[5-10)	[10-	Total (age)
[18-20)	996	439	367	163	87	87	36	2,172
[20-23)	2,286	975	806	443	211	218	171	5,110
[23-26)	1,996	985	866	456	205	241	230	4,979
[26-30)	1,591	954	839	437	215	256	320	4,612
[30-34)	885	563	541	340	164	237	302	3,032
[34-38)	461	318	397	309	164	351	114	2,114
[38-	633	478	692	525	300	340	20	2,988
Total (duration)	8,848	4,712	4,508	2,673	1,346	1,727	1,193	25,007
Number of observations across tenure and age groups. "Age" refers to the workers age at the time of recruitment. Sample consists of white-male U.S. workers. Data: <i>NLSY 79</i>								

Table 2.5: Summary Statistics I

The resulting data set consists of 25,007 observations of employment spells, 90.61% of which are uncensored. Table (2.5) depicts in detail how the total number of observations is distributed among the different tenure and age groups. Notice again that the term "age" refers to the worker's age at the time when the match was created and not to the worker's current age. Not surprisingly, the main bulk of employment spells in the data set are of a short duration. The proportion of jobs surviving into higher tenure classes shrinks as duration increases. Hence, the distribution of jobs across duration is strongly left skewed. This pattern holds throughout all age cohorts. Notice, however, that cohorts with higher age generally have a larger proportion of medium- and long-term jobs than young cohorts.

Table (2.6) provides summary statistics about the worker characteristics. The first two rows in the upper panel depict the mean and the standard deviation of workers' age across tenure classes. From this we observe that there is a positive relationship between the duration of an employment spell and the average age at which an individual has started working in the job. In short, matches which last longer are, on average, created by higher age co-

horts. Short-term jobs, i.e. those that end within half a year, are populated by workers aged on average 25.78 years whereas long-term jobs, those that last up to 10 years, are created by workers aged 30.32 years on average. However, the positive relation that we observe is not particularly strong. The standard deviation of age is generally quite large across tenure classes.

Job duration	[0-0.5)	[0.5-1)	[1-2)	[2-3.5)	[3.5-5)	[5-10)
Mean age	25.78	27.01	28.32	29.37	29.78	30.32
(Std.dev. Age)	(6.28)	(6.81)	(7.57)	(7.65)	(7.71)	(6.98)
Mean ω	-0.105	-0.032	0.043	0.079	0.127	0.136
(Std.dev. ω)	(0.35)	(0.42)	(0.48)	(0.51)	(0.57)	(0.56)

Initial age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)
Mean tenure	1.435	1.821	2.082	2.444	2.913	3.003
(Std.dev tenure)	(2.87)	(3.85)	(4.05)	(4.21)	(4.18)	(3.29)
Mean ω	0.00	0.00	0.00	0.00	0.00	0.00
(Std.dev. ω)	(0.30)	(0.33)	(0.39)	(0.48)	(0.51)	(0.59)

Upper panel: Mean and the standard deviation of workers' age across tenure classes (first two rows). Mean and the standard deviation of ω across tenure classes (third and fourth row). Lower panel: Mean job duration across age groups (fifth row). The standard deviation of ω (last row). Sample consists of white-male U.S. workers. Data: *NLSY* 79

Table 2.6: Summary Statistics II

The next two rows in the upper panel depict the mean and the standard deviation of ω across tenure classes. The observations here are in line with our theory. A central prediction of our model is that matches which pay above-average initial wages are likely to survive longer. This relation can also be found in the data. Short term jobs had payed initial wages which were below the average by about 10.5%. On the other hand, initial wages in medium and long term jobs were roughly 13% above average. Clearly, the survival probability of a match is positively related to the initial wage that is payed. According to our model, this relation originates in agents' beliefs that the match quality is a good predictor of the likelihood of survival. This in turn determines the value of a job and thereby affects the remuneration payed to the worker.

The lower panel in Table (2.6) reports summary statistics for different age cohorts. The first row contains the mean duration of an employment spell across age cohorts. We observe, basically, a mirror image of the first row in the upper panel. In short, the average duration of matches formed by young cohorts is substantially shorter than for jobs created by experienced workers. The average job duration more than doubles as we move from workers aged 18 – 20 years to workers aged 34 – 38 years. Not surprisingly, the mean value of ω is zero. This is by construction, as we measure ω as the deviation of an individual’s wage from the **average** wage in the **same** age cohort.

Arguably more interesting is the last row, i.e. the standard deviation of ω . We see a substantial increase in the standard deviation of ω , implying that there is more dispersion in initial wages for experienced workers than young workers. The dispersion roughly doubles as we move from the lower to the top age interval. This relation also holds when we use individuals’ labor market experience instead of their age as the criterion. The observed increase in the dispersion can be explained by the stochastic nature of labor market transitions. Within a group of initially similar workers, the degree of heterogeneity naturally rises over time as individuals build up their own labor market histories. More heterogeneity across individuals’ characteristics will be reflected by more dispersed initial wages.

2.8 Estimation Results

2.8.1 Coefficient estimates of the auxiliary model

In this section we report the coefficient estimates of the auxiliary model using the observed data. Before we bring the model to the data, we need to specify a grid for the job durations, τ_n , and age categories, a_k . Clearly, the distribution of job separations across duration is strongly skewed to the left as most separations occur early in a firm/worker relationship. To account for that we choose a finer grid for short durations and coarser one for long durations. Choosing a grid that is too fine, however, has the disadvantage that relatively few observations fall in each interval leading to imprecise estimates of β .

Bearing in mind these considerations we specify the following grid $\{0, 1, 2, 3, 5, 7, 10\}$ ²³. The choice of the age categories follows a similar logic²⁴. As documented by Section 2.2, most

²³Job duration is measured in years.

²⁴Recall that age in this framework refers to the worker’s age at the time of the match formation and not

of the labor market transitions, occur in the early years of an individual's career. To account for that, we make the age categories narrower for young individuals and wider for more experienced workers. In particular we choose the following grid for age $\{18, 20, 23, 26, 30, 34, 38, \}$. The coefficients to be estimated are

$$\beta = \begin{pmatrix} \eta_{18|20} & \eta_{20|23} & \eta_{23|26} & \eta_{26|30} & \eta_{30|34} & \eta_{34|38} & \eta_{\omega}, \\ \lambda_{0|1} & \lambda_{1|2} & \lambda_{2|3} & \lambda_{3|5} & \lambda_{5|7} & \lambda_{7|10} & \end{pmatrix}$$

The maximum likelihood estimator of β is given by the solution to the maximization problem stated in (3.4). Table (2.7) reports the results. Asymptotic standard errors are reported in parentheses. The coefficients of the baseline hazard rates, λ_i , are all very precisely estimated and the values are in line with expectations. The probability of separating from an employer is decreasing in tenure. This is a standard prediction of models in which firms and workers learn about the match quality. It is interesting to note though that the decrease is not linear in tenure.

At the beginning of an employment relationship the baseline hazard rate is fairly high but, conditional on staying in the job, it falls rather rapidly and subsequently levels out at a rate that is substantially lower than the initial value. This pattern is intuitive since most of the learning about the match quality takes place within the first few periods of an employment relationship. Matches that are revealed to be good continue to exist whereas bad matches break up. Thus, after an initial period of learning and selection, each surviving match is less likely to be of bad quality and leads to the non-linear decline in the baseline hazard rate. I also report the average hazard rate for each job duration. It is computed as the mean hazard rate of all observed individuals using the estimated coefficients. For instance, the mean hazard rate in the population of separating from an employer in the tenure interval $[\tau_{i-1}, \tau_i)$ (conditional on still being in the job) is given by $\frac{1}{J}\lambda(i) \sum_{j=1}^J \exp(\eta^j + \eta_{\omega}\omega_j)$.

Turning to the estimates of the age coefficients, η . Recall that each of the coefficients is linked to a dummy variable that is unity for the age interval that corresponds to the individual's age at match formation, and zero for all other age intervals. Furthermore, recall that the "missing" age cohort, which serves as the means of comparison, is given by the cohort with age ≥ 38 years. Therefore, any of the age coefficients reported in Table (2.7) has to be interpreted as the difference in the hazard rate between age groups k and $K + 1$, everything else being equal.

at the time of separation.

In addition to the coefficient estimates I also compute the marginal age effects in the population, the results of which are reported in the column on the far right of Table (2.7). Both, the estimates and the marginal effects suggest that the age at match formation is a fundamental determinant of an individual's probability of separating from an employer.

Coefficient	Estimate	A.H. (%)	Coefficient	Estimate	M.E. (%)
$\lambda_{0 1}$	0.4925 (0.0133)	0.885	$\eta_{18 20}$	0.7992 (0.0345)	122.4
$\lambda_{1 2}$	0.3187 (0.0095)	0.506	$\eta_{20 23}$	0.6997 (0.0306)	101.3
$\lambda_{2 3}$	0.2249 (0.0077)	0.339	$\eta_{23 26}$	0.6089 (0.0308)	83.85
$\lambda_{3 5}$	0.1638 (0.0058)	0.242	$\eta_{26 30}$	0.4653 (0.0320)	59.24
$\lambda_{5 7}$	0.1090 (0.0050)	0.153	$\eta_{30 34}$	0.2664 (0.0336)	30.53
$\lambda_{7 10}$	0.0876 (0.0044)	0.122	$\eta_{34 38}$	0.0814 (0.0384)	8.480
			η_{ω}	-0.498 (0.0178)	-0.497

$\lambda_{\tau|\tau'}$ denotes the estimate of the baseline hazard rate for the interval $[\tau, \tau')$, $\eta_{a|a'}$ is the coefficient estimate for the age cohort $[a, a')$, η_{ω} is the coefficient estimate for the variable ω , A.H. denotes the average effect and M.E. is the marginal effect. Standard errors are in parentheses.

Table 2.7: Estimation results: Auxiliary model using observed data

A worker that starts a new job at the age of 18 – 20 is more than twice (122.39% to be precise) as likely to separate from the current employer than a worker that starts a new job at (or above) the age of 38. This effect, however, declines dramatically with age, particularly within the first 10 years after labor market entry. An individual that starts a new job at the age of 30 is already about 60% less likely to separate from the current employer than they were 10 years previously. For ages higher than 34 years the effect seems to vanish. The

marginal effect of age in the cohort 34 – 38 years is only 8.48%, implying that a job that is started at this age is almost as likely to end than a job that is started at age 38 years or higher.

According to the structural model, this effect on the separation probability is due to imperfect information in the hiring process. These imperfections are linked to an individual's labor market experience since for young, as opposed to experienced, workers there is less information available to facilitate the assessment of the worker's productive ability. Over time, as an individual stays in the labor market, more and more relevant information accumulates and so the effect on the separation probability becomes less and less important. In the data this is reflected by a decreasing age coefficient. Moreover, the pattern in Table (2.7) suggests that the process which leads to a reduction in the importance of the age effect is strongly concentrated in the first 10 years of an individual's entry into the labor market. At later stages of the career, particularly from age 34 years onward, it becomes negligible.

The lower panel of Table (2.7) depicts the estimate of the coefficient of ω . The estimate is highly significant and has a negative sign. I also report the marginal effect, which is computed as the mean percentage change in the hazard rate in the population induced by a +1% gap of individual's initial wage to the mean initial wage of the respective age cohort. The value of the point estimate implies that a positive 1% deviation of a worker's initial wage from the mean initial wage in the same age cohort leads to a reduction in the individual hazard by roughly 0.5%. Regarding the sign of η_ω , the structural model yields the same prediction.

This is based on the following mechanism. When forming a match agents observe a noisy signal about the unobserved quality of the match. This signal is used to draw inference about the true quality, resulting in a certain belief. Agents are rational in the sense that their beliefs are, on average, consistent with the true match quality. Good matches persist, bad matches break up. Consequently, the value of the initial beliefs is indicative for the probability of separating from an employer. For this reason they also determine the joint surplus of a job and thereby also the wage that is paid to the worker.

The hazard rates in Table (2.7) are computed across all age cohorts. Figure (2.3) reports the hazard rates for selected age cohort separately²⁵. Two things are worth noting at this

²⁵Table (2.14) in Appendix F contains the corresponding numbers for all the age groups we consider in the estimation.

point: (1) At low job durations, workers of all ages feature a high probability of separating from the employer. Though the probability is substantially higher for young workers than for more experienced ones. (2) Throughout all tenure classes the hazard rate for young workers is approximately twice as high as that for experienced workers. Therefore, the percentage decline with tenure is roughly the same across cohorts, around 84%. The relative difference is constant, whereas the absolute difference in the hazard rates between young and experienced workers shrinks dramatically, from 0.5920 for newly created jobs to 0.0967 for jobs with 7 – 10 years of tenure. Thus, it is clearly the case that newly created jobs are much more likely to break up when a young worker is involved. Interestingly, for established jobs it matters much less if a young or an experienced worker is involved, as the gap in the separation probability narrows considerably.

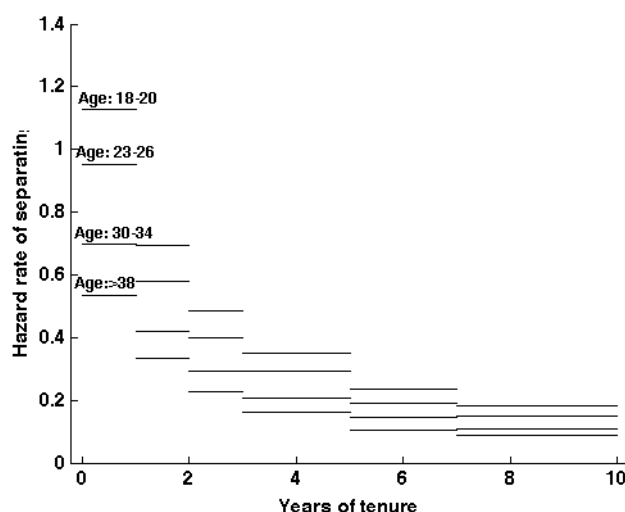


Figure 2.3

Actual hazard rate of separating from a job for selected age and tenure groups. Source: Own calculations based on data from the *NLSY* 79.

In order to assess the aptitude of the auxiliary model for capturing the main statistical properties of the data I perform a series of hypotheses tests. Checking the standard errors of the individual coefficient estimates can be considered as a first test to evaluate the explanatory power of the model. It appears that each of the estimates is statistically different from zero at the 98% confidence level²⁶.

²⁶In a preliminary step, I also included a worker's total labor market experience (measured as the worker's cumulated number of weeks being employed since labor market entry) as an explanatory variable. This

First, I perform a likelihood ratio test of the hypothesis that the age coefficients are all the same, i.e. $\eta_{18|20} = \eta_{20|23} = \dots = \eta_{34|38}$. The value is 446.8 with 5 degrees of freedom which rejects the hypotheses at the 99% confidence level. At the same level of confidence, I reject the hypotheses that the baseline hazard rates are all constant across tenure intervals, i.e. $\lambda_{0|1} = \lambda_{1|2} = \dots = \lambda_{7|10}$. Next, I test for the hypothesis that the hazard rates for any two consecutive intervals are the same, i.e. $\lambda_{\tau|\tau'} = \lambda_{\tau'|\tau''}$. The hypotheses $\lambda_{0|1} = \lambda_{1|2}$, $\lambda_{1|2} = \lambda_{2|3}$, $\lambda_{2|3} = \lambda_{3|5}$, $\lambda_{3|5} = \lambda_{5|7}$ are rejected at the 99% level, whereas $\lambda_{5|7} = \lambda_{7|10}$ is rejected at the 97.5% level. Next, I perform a similar test to check whether the age coefficients of two consecutive age cohorts are constant, i.e. $\eta_{a|a'} = \eta_{a'|a''}$. This test rejects $\eta_{18|20} = \eta_{20|23}$ at the 97.5% confidence level, whereas the remaining hypotheses, $\eta_{20|23} = \eta_{23|26}$, $\eta_{23|26} = \eta_{26|30}$, $\eta_{26|30} = \eta_{30|34}$, $\eta_{30|34} = \eta_{34|38}$, are all rejected at the 99% level.

Lastly, I test whether the choice of the age grid was an appropriate one. To this end, I first estimate the auxiliary model using a finer grid and then use the associated value of the likelihood function to perform a likelihood ratio test for the hypothesis that the grid in the original specification is as good as the finer one. Here, the term "good" refers to the difference in the likelihood of both specifications. When using the grid $\{18, 19, 20, 21, 23, 26, 28, 30, 32, 34, 38\}$ I can reject the hypotheses that the original grid, $\{18, 20, 23, 26, 30, 34, 38, \}$, would perform as well at the 75% confidence level only. Extending the grid to include higher ages, i.e. $\{18, 19, \dots, 24, 26, \dots, 44\}$, leads to the same result.

The results of these tests demonstrate that the chosen auxiliary model provides a good statistical description of the underlying data along the dimensions that are needed for the identification of the model's structural parameters. These dimensions are meant to capture any age-cohort effects that, according to the model, are causal for differences in the job separation and job turnover behavior of (a) individuals of different age and (b) along the working life cycle of each individual.

measure, however, turned out to be highly correlated with the individual's age variable, causing problems of multicollinearity.

2.8.2 Estimating the parameters of the structural model

Estimating the structural parameters means finding a $\tilde{\phi}_S^H$ that solves the minimization problem stated in (2.20). Performing one evaluation of the objective function involves the following steps:

1. Given the structural parameters ϕ the model is solved for the optimal policy functions.
2. With those at hand we can simulate the life path of a number of individuals to obtain data on the individuals' employment spells.
3. Using the simulated data we then estimate the parameters of the auxiliary model and finally evaluate the quadratic form in (2.20).

In order to solve the model we need to take a stand on a variety of issues. Concerning an individual's life-cycle, we need to specify (a) the survival probabilities ρ_k , for $k = \{1, 2, \dots, \bar{K}\}$ and (b) the number of periods an individual stays in the labor market, set \bar{K} . The data that is used here does not allow for the identification of death. We do observe individuals dropping out of the survey but the cause of dropping out is not specified. It is unlikely, though, that death plays an important role. The individuals observed in 2006 (the year of the last survey) are aged 42 – 48 years. According to the U.S. National Center for Health Statistics, those individuals can expect to have an additional 30.5 – 36.9 years to live, depending on their age within the cohort. Moreover, the annualized probability of death for this age cohort is less than 0.25% indicating that, for those individuals, the risk of death is fairly low. In light of this, we set $\rho_k = 1$ for $k = 1, \dots, \bar{K}$. Furthermore, the time dimension of the *NLSY79* survey is still limited in the sense that we do not observe individuals up until retirement. Consequently, we can not compute \bar{K} from the survey data. Instead we set $\bar{K} = 45$, implying that individuals spend at maximum 45 years in the labor force.

In the simulation step we simulate the working life cycle of $\mathcal{S} = 10^5$ individuals. For each individual we observe a certain number of employment spells, the number and the length of which depend on the set of parameters ϕ ²⁷. Therefore, the number of observations in the estimation is endogenous and dependent on the set of parameters. This is undesirable since a low number of observations implies relatively inaccurate estimates of the structural parameters. To circumvent this problem I introduce a penalization term into the algorithm.

²⁷If, for instance, the policy functions implied by ϕ lead to high threshold beliefs then few individuals will accept job offers and the number of employment spells in the resulting data set will be relatively low.

In particular, the algorithm penalizes those parameter constellations for which the implied number of observations is lower than a certain critical value. As a critical value I use $2 * \mathcal{S}$. Furthermore, there are assumed to be two types of worker, denoted by $i \in \{i_1, i_2\}$. Before bringing the structural model to the data we need to make the following adjustment. Notice that the variance of the noise component is the only characteristic that distinguishes young from experienced workers. Therefore, the model would attribute all the differences in the job separation across age entirely to the information imperfections. This is clearly an overstatement of the importance of the information channel. Estimates based on the current specification would most likely be biased. Obviously, there are other sources that cause separation rates to differ across age. To account for that we do not ex-ante restrict (a) the exogenous rate of separation σ and (b) the job quality, governed by α , to be the same across types. Instead, we allow σ and α to vary across i -types, i.e. we consider $(\sigma_{i_1}, \sigma_{i_2})$, $(\alpha_{i_1}, \alpha_{i_2})$ and estimate them separately.

Moreover, I restrict the process that governs changes in workers' types to be irreversible. In particular, I set $\mu_{2|2} = 1$, $\mu_{2|1} = 0$, which excludes the possibility of type "depreciation". This assumption does not seem to be very restrictive. A worker's type is meant to capture factors that affect the ability to screen a worker, such as the stock of information about the worker's labor market history. The more information, the better an applicant can be screened and categorized by the hiring firm. Such information does not get lost over time. Therefore, it is arguably a minor point to neglect type depreciation. Therefore, the only free parameter that governs changes in workers' types is $\mu_{1|2}$. Notice that $\mu_{1|1}$ can be recovered using $\mu_{1|1} = 1 - \mu_{1|2}$, once we know $\mu_{1|2}$. In total the model consists of 14 structural parameters. These are:

$$\left(\beta, \tilde{\lambda}, e, y^g, y^b, b, \mu_{1|2}, \varphi, \sigma_{i_1}, \sigma_{i_2}, \alpha_{i_1}, \alpha_{i_2}, s_{i_1}, s_{i_2} \right)$$

Ten of these parameters are estimated, whereas the remaining four, $(\beta, \tilde{\lambda}, e, y^g)$, are set beforehand. As a normalization, we set $y^g = 1$, due to the fact that, in terms of the expected per-period output of a match, the important consideration is not the level of y^g but the difference $y^g - y^b$. Therefore, we fix y^g and estimate y^b . We set the personal discount factor β equal 0.9569 which implies an annual interest rate of 4.5%. Typically, it is difficult to get reasonable estimates for the personal discount factor β , especially in this context where the existence of individuals' heterogeneity and a finite life cycle render the estimation β within meaningful bounds even more problematic. We set the firm's bargaining power e equal to 0.5, which is a standard value used in the search and matching literature. The estimation strategy used here is unlikely to properly identify the arrival rate of new job offers, $\tilde{\lambda}$. This

parameter is a key determinant of the duration of an unemployment spell but it has little effect on the duration of an employment spell. The hazard rate approach which we adopted focuses on the duration of jobs but it does not process any information on individuals' unemployment spells. Consequently, $\tilde{\lambda}$ would be poorly identified and any estimate of it would be meaningless. We set $\tilde{\lambda} = 1$, which we believe is a reasonable choice given that we use yearly observations and the average unemployment spell in the *U.S.* is significantly shorter than a year.

To estimate the parameter vector $\tilde{\phi}_S^H = (y^b, b, \mu_{1|2}, \varphi, \sigma_{i_1}, \sigma_{i_2}, \alpha_{i_1}, \alpha_{i_2}, s_{i_1}, s_{i_2})$ I use a hybrid method as proposed by Nagypal (2007). This method first evaluates the objective function along a relatively coarse grid. The initial grid that I chose has dimension $4 \times 2 \times 6 \times 4 \times 3 \times 3 \times 4 \times 4 \times 3 \times 3$. To economize on computational time I start out with a low number of simulations. In each subsequent round, I increase the number of simulations leading to more precise evaluations of the objective function. However, I consider only those grid points that perform well in the previous round. All the remaining points are eliminated. After this refinement I am left with a grid with 279 points. The values of the objective function on this grid range from 6.46×10^{-6} to 2.15×10^{-5} . The method is described in detail in Appendix E of Nagypal (2007)²⁸. In the next step, I use the remaining grid points as starting values in a simplex algorithm that solves the minimization problem in (2.20). Table (2.8) reports the estimates of the structural parameters. To test whether the structural model is well specified I perform a global specification test that was proposed by Gourieroux et al. (1993). This test is based on the minimized value of the objective function with the statistic

$$\zeta_S = \frac{SH}{1+H} \times \min_{\phi} \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^H \beta^h(\mathbf{y}_S, \phi) \right) \mathcal{Q}^* \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^H \beta^h(\mathbf{y}_S, \phi) \right)' \quad (2.23)$$

that is asymptotically distributed as a χ^2 with $\dim(\beta) - \dim(\phi)$ degrees of freedom. The test statistic in our case is equal to 1.374, which is well below the critical value of 7.82 for the 95% confidence level with 3 degrees of freedom. Therefore, we can not reject the null hypotheses that the model is well specified. The standard errors depicted in Table 2.8 are computed by evaluating the expression in (2.21). For that we consider a 0.0001% deviation from the estimates in Table 2.8. To get accurately estimated standard errors we rely on a very large set of simulated data. In particular, we simulate the working life cycles of 20 million individuals, which results in approximately 120 million observations of employment

²⁸Using the terminology by Nagypal (2007), I use a total of 4 rounds in which the number of simulations is $\{5 \times 10^3, 10^4, 5 \times 10^4, 10^5\}$.

spells. The standard errors in Table 2.8 suggest that the coefficients are estimated fairly precisely. This does not, per se, rule out the possibility that the minimum found is the local minimum.

To double-check the accuracy of the estimation procedure and in particular, to identify flat regions of the objective function surface we compute the objective function fixing all but one parameter at the estimated value, varying the remaining parameter over a range of feasible values. The results are depicted in Figure 2.8 in Appendix F. Most importantly, we do not find any notable flat regions of the criterion function, which confirms the accuracy of our estimates²⁹. The possibility of being trapped in a local minimum is nearly entirely ruled out by the operating mode of the grid search algorithm that was used to obtain the starting values. This algorithm provides us with a range of values, each of which is potentially close to a (local) minimum. Using many different starting values in the subsequent simplex-search algorithm guarantees that we can identify (and dismiss) local minima by comparing the values of the criterion function at all of the points at which the algorithm terminates the search process.

The estimate of $\mu_{1|2}$ is equal to 0.4134, meaning that an employed type 1 worker faces a 41.34% (annualized) probability of becoming a type 2 worker, conditional on staying employed. Consequently, a young worker would need, on average, 2.42 years of tenure on a single job before receiving the type-upgrade. However, jobs often do not survive that long and employment spells are typically interrupted by periods of unemployment. Particularly, the high turnover for young cohorts is detrimental to a quick upgrade as it prevents the accumulation of years of tenure on a given job. Thus, it takes substantially longer than 2.42 years until a type change occurs. In the simulated version of the model (an explanation

²⁹An interesting case emerges when we vary s_2 . The panel in the last row in Figure 2.8 illustrates that for values of $s_2 > 0.26$ the criterion function increases markedly. This is due to the following reason. s_2 is defined as the standard deviation of a truncated normal distribution with domain $[0, 1]$. To draw random numbers that follow a truncated normal distribution we have to compute the standard deviation (together with the mean) of the corresponding un-truncated normal. The issue is, however, that for a truncated normal with mean in $[0, 1]$ there exists an upper bound on the standard deviation which is around 0.288. Any arbitrarily high value for the standard deviation of the un-truncated normal translates into a value for s that is always below this upper bound. Consequently, for values of s_2 close to that upper bound the implied random draws for agents' beliefs are that much dispersed so that, in the end, very few draws fall above the threshold belief. This means that very few individuals get matched and we observe very few employment spells in the resulting simulated data. As mentioned previously, we incorporate a penalty term into the algorithm that adds a very high positive number to the objective function whenever the number of simulated employment spells falls short of a critical value. The case of s_2 being close to the upper bound is such a case, hence we observe the criterion function jumping upwards.

follows), where we take into account all the labor market transitions of an individual, we find that the average worker moves from type 1 to type 2, 5.25 years after labor market entry. The quality of a match is initially unknown but firms learn about it, over time, on the basis of observed output. In each period there is a constant probability, estimated to be $\varphi = 0.6157$, that the match quality is fully revealed. As a result, it takes, on average, 1.62 years until the firm learns the true quality of the match.

Parameter	Explanation	Estimate	
y^b	Output of a bad firm/worker match	-3.2433	(0.3350)
b	Flow value of unemployment	0.4333	(0.1233)
$\mu_{1 2}$	Probability of a type change	0.4134	(0.0726)
φ	Probability that the match quality is revealed	0.6157	(0.0792)

Parameter	Explanation	Estimate	
		Type 1	Type 2
σ_i	Probability of exogenous job separation	0.2497 (0.0122)	0.1084 (0.0323)
α_i	Mean of agents' beliefs	0.3781 (0.0969)	0.7571 (0.0706)
s_i	Standard deviation of agents' beliefs	0.0962 (0.0097)	0.2030 (0.0156)

Estimates of the structural parameters of the model presented in Section 2.3. The standard errors are reported in parentheses.			
--	--	--	--

Table 2.8: Estimation results: Structural Model

The lower panel of Table (2.8) reports the estimates of the parameters that are allowed to differ across i types. In our setup, σ is the rate at which matches break up for exogenous reasons. In the estimation it can be thought of as a residual that captures all the separations

that are not due bad match quality. The (annualized) rate of exogenous match separation for matches with type 1 and type 2 workers is estimated to be 0.2497 and 0.1084, respectively. The value for experienced workers is close to that usually reported by empirical studies. Davis and Haltiwanger (1992), for instance, finds that in the U.S. the rate of exogenous job destruction is 11.3% per year. The rate for young cohorts, we find, is substantially higher than the rate for experienced worker. These numbers imply that matches of experienced workers are hit by an exogenous shock, which leads to a break-up, after an average of 9.2 years, whereas matches of young workers are hit after about 4 years. To test whether the estimates of σ_i are statistically different from one another I perform a Wald test of the hypothesis that $\sigma_1 = \sigma_2$, i.e. the rate of exogenous separation is the same across types. The statistic associated with null hypotheses is given by

$$\zeta^W = c(\tilde{\phi}_S^H) (CWC')^{-1} c(\tilde{\phi}_S^H)$$

where $\zeta^W \sim^a \chi_2^2$ and \mathcal{W} is the asymptotic variance-covariance matrix determined in (2.21). $c : \Phi \rightarrow \Re$ is a continuously differentiable function on the parameter space Φ that incorporates the null hypotheses: $H_0 : c(\phi_0) = 0$ and $C(\tilde{\phi}_S^H)$ is the gradient of $c(\cdot)$ evaluated at the optimal unconstrained estimator $\tilde{\phi}_S^H$. The test statistics associated with the null hypothesis is 10.08, which is above the critical value of 9.2 for the 99% confidence level with 2 degrees of freedom. Therefore, we reject the null hypotheses that the exogenous rate of job separation is the same for type 1 and type 2 workers.

In our model α has a threefold interpretation. On one hand, it is the mean of the distribution from which the true, but unknown, job quality π is drawn from. Given the zero-mean assumption for the noise component ϵ , α also represents the mean of the observed signal, γ , and provided how agents form their beliefs about π it is also the mean of the beliefs that are realized. The estimates we obtain for α_1 and α_2 imply that the true job quality is substantially lower for young workers. Notice, however, that α is not the average quality of all *realized* matches but of all *potential* matches. Not all matches are formed in equilibrium as encounters with a signal below the threshold do not result in the match formation. Hence, the actual average quality of existing match is higher and is equal to 0.524 for type 1 workers and 0.894 for type 2 workers for our purposes. This says that almost half of all employed type 1 workers, but just 10% of type 2 workers, are in a mismatch³⁰.

³⁰Notice that we are referring to matches whose quality is not yet revealed.

The large difference between the two groups can be explained by the type-specific noise component. Recall that young workers emit a less informative signal than experienced worker. A poor signal that is very noisy is more likely to indicate a good match than the same signal that is precise. Therefore, firms can risk hiring young workers with relatively poor signals as there is a substantial likelihood that the resulting matches are in fact good. Consequently, the hiring threshold is lower for type 1 workers and so is the average quality of existing matches. Lastly, I perform a Wald test on the hypothesis $H_0 : \alpha_1 = \alpha_2$ to check whether the difference in the estimated values is statistically significant. The test statistic is 13.09 and so we may reject the hypotheses that the job quality is the same across worker types at the 99% confidence level.

The goal of this paper is to assess the importance of worker-specific informational frictions for explaining the observed differences in labor market outcomes across age cohorts. In our model those frictions are captured by a noise component that distorts workers' signals. The worker specificity shows up in the variance of the noise component $\sigma_{\epsilon,i}^2$, which we assume to be different across worker types. $\sigma_{\epsilon,i}^2$ determines the variance of agents' beliefs via $s_i^2 = \sigma_\pi^4 / (\sigma_\pi^2 + \sigma_{\epsilon,i}^2)$. We did not estimate $\sigma_{\epsilon,i}^2$ itself but we can back it out from the expression $s_i = \sigma_\pi^2 / \sqrt{\sigma_\pi^2 + \sigma_{\epsilon,i}^2}$ using the estimated values for the standard deviation of agents' beliefs s_i . The estimates of s_1 and s_2 are respectively 0.0962 and 0.2030. We use a range of values for σ_π^2 and find that the ratio $\sigma_{\epsilon,1}/\sigma_{\epsilon,2}$ takes values of around 2.1. The implied signal to noise ratio is about 4.5 times higher for experienced workers than for young individuals. Lastly, I perform a Wald test to check whether the difference in the precision of workers' signals is significant. We find that the test statistics associated with $H_0 : s_1 = s_2$ is 12.78 and so we reject the hypotheses that the informational friction is **not** worker-type-specific.

2.8.3 Discussion of the Results

In this section, we assess how well the predictions obtained from the structural model match up with actual data. To that end, we use the estimated parameters and simulate the model to generate data on individuals' working life cycles. From the simulated data we compute a variety of statistics we can then use to compare to their counterparts in the observed data. First, however, we discuss the optimal policy functions implied by the model. Recall that the policy in this model is given by the threshold value of beliefs for both worker types. Figure (2.4) depicts the results.

As expected, the cutoff for both types is increasing in age. This is intuitive as an individual getting ever closer to the exit age becomes less and less valuable for a (new) match as the time horizon, over which a firm can collect the surplus, shrinks. To compensate firms for the shorter horizon and for the risk that the match might be bad, a higher surplus is required to make matching with older workers attractive. Consequently, the threshold increases. Notice, however, that the threshold is flat for almost the entire career of type 2 workers, as for these workers there is very little risk that the match turns out to be bad and so firms do not require any compensation.

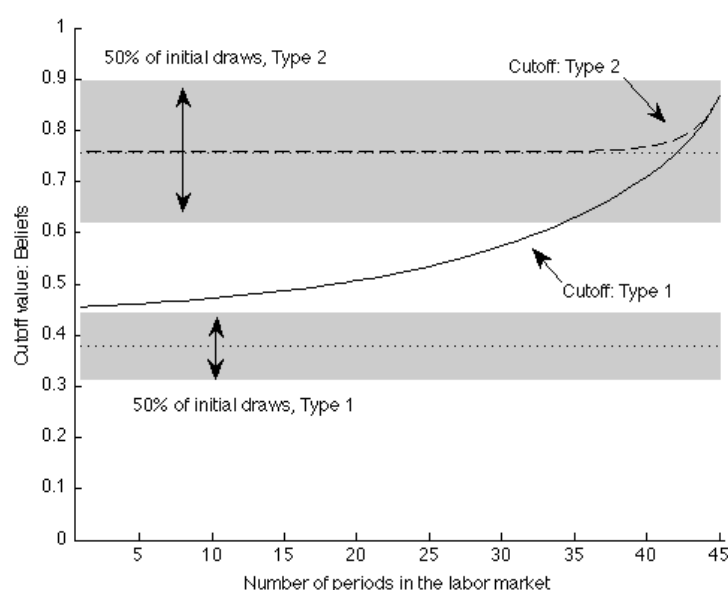


Figure 2.4

(a) Equilibrium threshold value of beliefs for type 1 and type 2 workers (solid and dashed line), (b) mean value of the initial beliefs (dotted lines), (c) 50% confidence bounds around the mean (shaded area).

We have demonstrated that the threshold for type 2 workers is higher than for type 1 workers in the simple case (see Section 2.4) and it also holds here. The reasoning follows as before. The outside option of employed type 2 workers is more valuable, as their expected value of a job is higher than that of a type 1 worker. Therefore, they can extract a higher fraction of the total surplus, which translates into an increased reservation belief. Figure 2.4 also contains the mean value of the initial beliefs (indicated by the dotted line) and the 50% "confidence bounds" around the mean (shaded area).

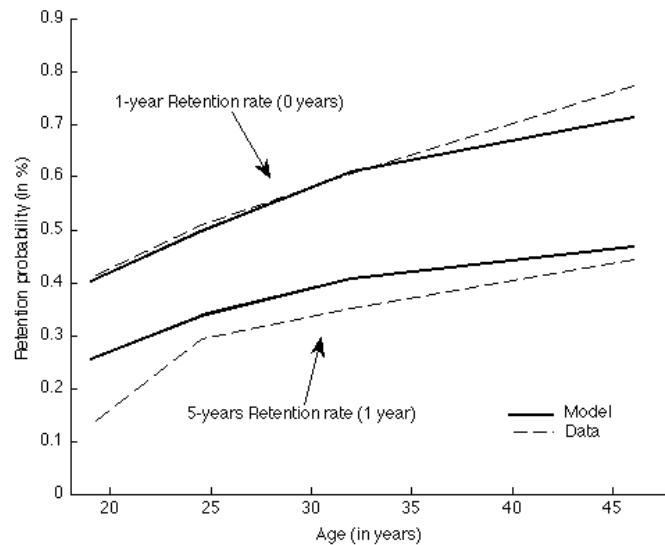


Figure 2.5

The 1-year retention rate for newly employed workers (upper two lines) and the 5-years retention rate for workers with one year of tenure (lower two lines). Comparing the results obtained from the model (solid line) with the counterpart in the data (dashed line). Source: Own calculations using data from the *NLSY 79*.

These bounds indicate the region in which 50% of all realized initial draws fall into. From there, we clearly see that the initial beliefs of type 1 workers are substantially less dispersed. This is a consequence of how beliefs are formed (see Equation (2.2)). More noise in the signal means that, when forming beliefs, more weight is put on the mean value and less on the actual signal. Thus, there is less dispersion in type 1 beliefs. Furthermore, notice that the cutoff value for type 1 is above the shaded area, meaning that type 1 workers are, on average, less successful in generating a signal that is sufficient for a match formation. Thus the noise in the signal (combined with a lower mean value of beliefs α_1) is basically a hiring barrier for those workers³¹.

³¹This has important implications for the duration of unemployment. According to the model, we should observe young (type 1) workers having longer spells of unemployment than experienced workers. In reality, however, this effect might be at least partly offset by a higher arrival rate of job offers for the young, as they are typically less specialized and so more willing to take on any given job. In our model, we abstract from this issue and cannot account for age-specific differences in the duration of unemployment.

One of the main findings in Section 2.2.1 was that young individuals are substantially more likely to separate from a new job than experienced workers. We compute job retention rates from the data and find that the likelihood of leaving a new job declines with age. Next we compute the same statistics using the simulated data. Figure (2.5) reports the results for the 1-year retention rate for newly employed workers (the upper two lines) and the 5-years retention rate for workers with one year of tenure (lower two lines)³². The solid line is the prediction of the model and the dashed line is the counterpart in the data.

Table (2.9) contains the results for the remaining retention classes. For all the cases, the model can capture very well the increase in the retention probability with age. Also the level that is predicted is broadly in line with the actual data, though for young workers, the model slightly overpredicts the short-horizon retention rate. For experienced workers the predictions of the model are generally closer to the empirical facts.

Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.41 (0.40)	0.51 (0.49)	0.60 (0.61)	0.77 (0.71)
2 years	0.21 (0.25)	0.33 (0.36)	0.42 (0.48)	0.61 (0.59)
3 years	0.12 (0.20)	0.26 (0.31)	0.33 (0.41)	0.45 (0.51)
4 years	0.09 (0.17)	0.21 (0.26)	0.27 (0.36)	0.41 (0.45)
5 years	0.07 (0.15)	0.17 (0.23)	0.24 (0.32)	0.36 (0.41)
Job retention rates for newly employed workers: Rows represent different retention periods, columns indicate workers' age at the time of recruitment. Numbers in bold (normal font) represent the model (data).				

Table 2.9

We also compute the retention probabilities for workers with one year of tenure. Table (2.14) in Appendix E reports those results. The pattern for that tenure class is generally very similar to that of newly employed workers. The previous results suggest that job attachments are rather fragile for young workers but become more and more durable in the later stages of a career. In this dimension, the model matches the empirical facts very well.

³²Recall that the x -years retention rate is the probability that a worker with a certain tenure keeps her current job for x more years.

An aggregate implication of unstable jobs is high job turnover among young cohorts. In the data this is clearly observable as illustrated in Section 2.2.1. In what follows, we present some tables and figures documenting how well our model can account for observed differences in job turnover across age cohorts. In a world with high turnover, we should observe short short employment spells and individuals holding many jobs in a given period of time. Table (2.10) reports, for different age cohorts, the fraction of workers within a given tenure class.

Tenure / Age	[18-20)	[23-26)	[30-34)	[38-
0 – 2	0.89 (0.84)	0.61 (0.39)	0.37 (0.23)	0.24 (0.21)
2 – 5	0.11 (0.16)	0.27 (0.39)	0.27 (0.26)	0.21 (0.22)
5 – 10	0.00 (0.00)	0.12 (0.22)	0.24 (0.31)	0.22 (0.24)
> 10	0.00 (0.00)	0.00 (0.00)	0.12 (0.20)	0.33 (0.33)
Proportion of workers with $\tau \in \{[0, 2), [2, 5), [5, 10), > 10\}$ years of tenure within a given age cohort. Numbers in bold (normal font) represent the model (data).				

Table 2.10

The numbers in bold represent the outcome of the model. Clearly, young workers are over-represented in short-term jobs as, for them, little time has passed since entry into the labor market. Then, as the career progresses, more and more workers move into medium and long term jobs. Actually, for the youngest cohort the model can replicate fairly well the actual tenure "distribution". It over-predicts the fraction of young workers that move into medium term jobs (second column). Also the results for more experienced workers are broadly in line with the empirical facts.

One of the key facts presented in Section 2.2.1 indicates that there is a lot of job-shopping among the young but substantially less turnover for experienced workers. As a consequence, individuals typically hold the vast majority of all their lifetime jobs within the first 10 years of the career. In Table (2.11) we report the average number of jobs a worker holds until a certain age as a fraction of the career total. The upper panel shows the data and the lower panel contains the numbers produced by the model.

Generally speaking, the model can replicate quite well the observed pattern, i.e. the steep increase in the number of jobs during the early years in the career and the flattening

out in later stages. However, as mentioned previously, the model slightly underestimates the amount of turnover for young workers. Therefore, the initial increase is less pronounced than in the data.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
Data							
	0.332	0.484	0.629	0.746	0.806	0.882	0.963
	(0.211)	(0.236)	(0.232)	(0.211)	(0.201)	(0.165)	(0.084)
Model							
	0.211	0.376	0.539	0.661	0.765	0.854	0.948
	(0.185)	(0.209)	(0.213)	(0.203)	(0.185)	(0.156)	(0.093)
Each entry represent the average number of jobs a worker holds until age a as a fraction of the total number of jobs hold in the entire career. The standard errors are in brackets.							

Table 2.11

However, there is substantial heterogeneity in the number of jobs individuals hold along the career. This is indicated by the large standard deviation. In order to get a better picture, Figure (2.6) plots the mean number of jobs together with the 50% confidence area around the mean. This area determines the region in which 50% of the individuals belonging to a given age cohort fall into. In this figure, we see that there is actually a lot of overlap between the model outcome and the actual data.

Next, we report the number of newly employed workers as a fraction of total employment. Panel (a) in Figure (2.7) compares the model outcome with the empirical counterpart. In the data we define a newly employed as a worker having at most 20 weeks of tenure. In this dimension the model does extremely well. Table (2.10) shows that it generally underestimates the proportion of workers with tenure less than 2 years but it is highly accurate predicting the very short run horizon. Lastly, we depict the age-specific unemployment rate, obtained from the model, and compare it to the data. Panel (b) in Figure 2.7 contains the respective plots.

Qualitatively, the model can capture the observed decline in unemployment with age. However, quantitatively, the model overestimates unemployment for basically all age groups.

This is not too surprising as the empirical strategy, employed to estimate the structural parameters of the model, only uses information on individuals' employment spells but does not make use of information regarding unemployment spells. Thus, there is no target in the auxiliary model that captures the observed pattern of unemployment.

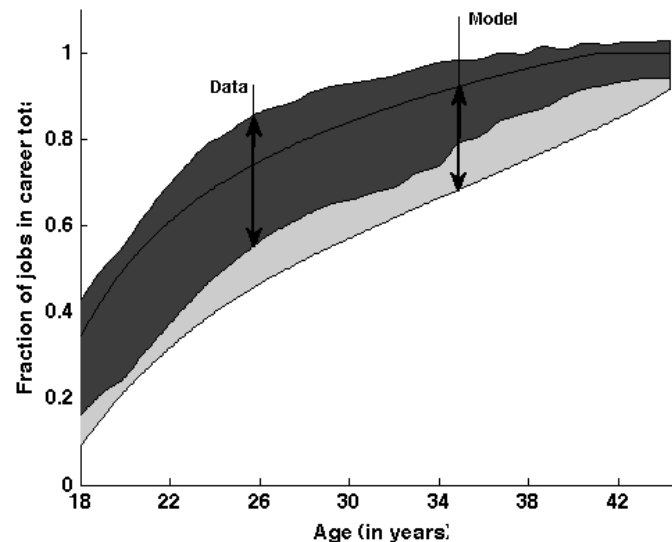


Figure 2.6

The 50% confidence area around the average number of jobs a worker holds until a certain age as a fraction of the career total. Comparing the model (light-grey area) with the data (dark-grey area). Source: Own calculations using data from the *NLSY 79*.

In the previous section, it can be seen that the model accounts fairly well for the observed age-specific differences in labor market outcomes. It was also demonstrated that all the parameters that were allowed to differ across *i*-types, i.e. σ , α and s , are actually statistically significant from one another. The question concerning the contribution of each of those parameters in explaining the observed life-cycle dynamics of individual job mobility remains. To address this question, we run an experiment in which we implement various restrictions on the parameters and compare the outcomes of the model obtained under the restrictions with both the data outcome and the results of the unrestricted model. For this, we take the results for the job retention probabilities as a standard of comparison. In particular, we use the 1, 3 and 5-year retention rates for newly employed workers.

The results are reported in Table 2.12. Each of the the first five columns depicts the difference in the $\{1, 3, 5\}$ -year retention rate between workers with a given age (top row) and workers aged 18 – 20 years. Basically, we capture the increase in job stability over the life cycle. The first row in each panel is the data outcome. The first entry, for instance, means that newly employed workers aged 23 – 26 years are 10.1% more likely to retain the job than newly employed workers aged 18 – 20 years. Each of the rows thereafter contain a particular outcome of the model obtained under a given parameter restriction.

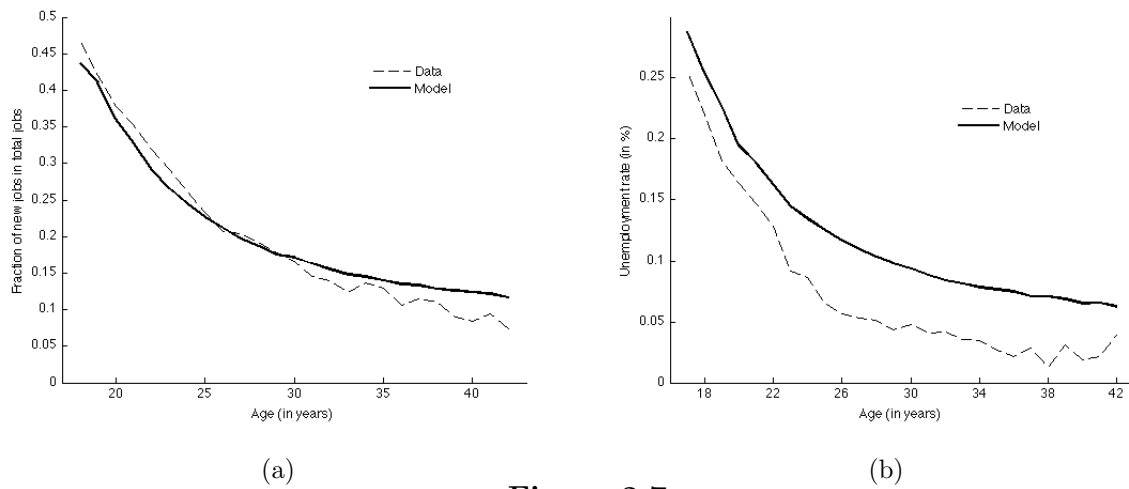


Figure 2.7

Panel (a): The number of newly employed workers with a given age as a fraction of total employment (of the same age group). Comparing the results obtained from the model (solid line) with the empirical counterpart (dashed line). Panel (b): Comparing the unemployment rate generated by the model (solid line) with the empirical counterpart (dashed line). Source: Own calculations using data from the *NLSY* 79

The restrictions we implement are as follows: (1) $\sigma_i = \bar{\sigma}, \alpha_i = \bar{\alpha}, s_i = \bar{s}, \forall i$ (second row), which means that we do not allow for any type-specificity of parameters. (2) $\sigma_i = \bar{\sigma}, \alpha_i = \bar{\alpha}$ (third row), which implies we allow only the noise component to be different across i types. (3) $\sigma_i = \bar{\sigma}$ (fourth row), so we allow α and s to differ. The last row depicts the outcomes of the unrestricted model that we obtain by using the parameter estimates from Table 2.8. In each of the restrictions, we fix the parameter at some value that is given by the arithmetic mean of the estimated, type specific values. For instance, $\bar{\alpha} = (\alpha_1 + \alpha_2)/2$. The leftmost column in Table 2.12 reports the average agreement of the respective scenario with the data-outcome. For brevity we only provide a description of the results for the 1-year job retention

rate, i.e. the first panel in Table 2.8. Notice that the unrestricted model (fifth row) can account, on average, for 98.5% of the increase in the observed 1-year rate.

Age	[23-26)	[26-30)	[30-34)	[34-38)	[38-	% explained
1-year job retention rate						
Data	0.101	0.139	0.198	0.245	0.364	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.026	0.055	0.092	0.122	0.141	0.418
α, s	0.077	0.135	0.192	0.233	0.272	0.873
σ, α, s	0.091	0.146	0.205	0.261	0.313	0.985
3-years rate						
Data	0.141	0.164	0.220	0.272	0.329	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.018	0.046	0.069	0.095	0.113	0.287
α, s	0.058	0.108	0.152	0.186	0.221	0.624
σ, α, s	0.095	0.155	0.212	0.265	0.316	0.906
5-years rate						
Data	0.096	0.118	0.169	0.189	0.289	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.014	0.028	0.046	0.064	0.076	0.259
α, s	0.039	0.072	0.104	0.124	0.152	0.562
σ, α, s	0.078	0.127	0.172	0.212	0.261	0.992
Each entry in the first five columns represent the difference in the $\{1, 3, 5\}$ -years retention rate between workers with a given age and workers aged 18 – 20 years. The first row in each panel is the data. The following rows represent the model outcome obtained when restricting certain type-specific parameters to be the same across i -types. We report the parameters that are allowed to differ across i -types. The leftmost column reports the average agreement with the data-outcome.						

Table 2.12

When we restrict all the parameters to be the same across i -types (second row) the model is not able to generate any life-cycle dynamics. The zeros indicate that, in this scenario, there isn't any change in the retention rate as workers age. Consequently, all of the age-specificity

remains unexplained. In the next case (third row), we still keep σ and α fixed but we allow s to differ across i -types.³³

Introducing worker-specific information imperfections considerably improves the empirical fit of the model. 41.8% of the observed increase in the 1-year retention probability can be accounted for solely by the worker-specificity of the imperfections. The contribution is lower but still substantial for the 3-year and 5-year retention rates. Keeping σ fixed but allowing, in addition to s , the job quality α to differ as well, adds another 45.5%. The remaining 11.2% can be explained by introducing worker-type-specific rates of exogenous match separation.

Clearly, the major part of the observed increase can be explained by introducing worker-type-specificity in α and s . Both parameters are fundamental for the signal extraction process, as α represents the mean of the observed signals and s relates to the amount of noise in a signal. The contribution of σ seems negligible, at least for the short-term horizon. However, when looking at the results depicted in the middle and lower panels of Table 2.12, it becomes clear that contribution of σ rises substantially when we consider the medium and long term horizon. For the 3-year and 5-year retention horizons the contribution of σ is 28.2% and 43%, respectively. There is an intuitive explanation for why σ is less important for explaining the short-run separation pattern but substantially more important for the medium and long run horizons.

In our model all the mismatches break up upon the revelation of the bad match quality. According to the estimates, the quality of a match is revealed, on average, after 1.62 years but an exogenous break-up occurs, on average, only after 4 years (young) and 9.2 years (experienced). Therefore, most of the match separation in the short run is due to endogenous separation whereas break-ups in the medium and long term are driven by exogenous shocks. Furthermore, the important factor for the magnitude of endogenous break-ups is primarily the distribution from which the signals are drawn. This distribution is shaped by α and s . As a result, it is α and s that govern the separation process in the short run whereas σ drives job destruction in the medium and long run.

³³In all the cases in which we allow parameters to differ across types we set them equal to the respective estimated value.

2.9 Conclusion

In this paper we present new empirical evidence documenting the life-cycle dynamics of individual job mobility in the U.S. labor market. Based on *NLSY 79* data, we find that the job retention probability for newly employed workers is strongly increasing with an individual's age and labor market experience. We estimate that the first 10 years of labor market experience raise the probability of retaining a new job for one year or longer by roughly 20%. Moreover, we find that, until the age of 45 years, a typical white male U.S. worker holds about 9 full time jobs, 50% of which are held within the first 5 years after labor market entry and roughly 75% within the first 10 years. This suggests an enormous amount of job turnover for individuals in the first years after labor market entry. In the later stages of a career, job attachments become substantially more durable and job changes less frequent.

To gather insights on the observed pattern of individual job mobility we construct and estimate a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. The key ingredient is that the imperfection is assumed to be worker-specific and in particular, it is linked to an individual's previous labor market history. We estimate the structural parameters of the model by indirect inference on data from the *NLSY 79*. Using the estimates we evaluate the empirical content of our framework by assessing how well certain predictions, regarding individual and aggregate labor market statistics, obtained from the structural model match up with actual data. We find that the model can capture very well the observed life-cycle dimension of a variety of individual labor market outcomes. In particular, it can account for the fact that job attachments tend to be very fragile early in an individual's career but become increasingly durable as workers accumulate more and more labor market experience. Furthermore, we find that the informational frictions considered in the model are key for replicating the life-cycle profile of individual labor market outcomes.

Beyond contributing to our understanding of the observed life-cycle dynamics of job mobility, the findings of this paper have important implications, especially for the design of optimal labor market policies. Our results stress that the accumulation of work experience is key for an individual's job stability and job security. With more experience accumulated on past jobs any future employment relationship will be more stable and secure. This implies a certain path dependency which is important for labor market policies in general. Any policy that deters young workers from moving into jobs, or that prevents firms from hiring them,

can be detrimental to the labor market prospects for an individual's entire career. Firing taxes or excessive employment protection are examples of policies that typically discourage firms from hiring young workers.

On the other hand, policies that foster the attachment of young workers to, and their integration into, the labor market can have significant positive employment and welfare effects as the work experience accumulated in the early years positively affects the job situation later in an individual's career. On-the-job training programs, such as the German apprenticeship program, is a good example for such a policy. The design of optimal labor market policies has to take a fundamental account of this life-cycle dimension. The aggregate welfare effects of a policy must be evaluated considering not only that, at each point in time, the policy can have a different impact on young individuals than on experienced workers but also that each individual is affected differently by the policy depending on the stage of the life-cycle.

Bibliography

- [1] Altonji, J.G., and Pierret, C.R., 2001. Employer learning and statistical discrimination, *Quarterly Journal of Economics*, 116 (February), pp. 313-350
- [2] Booth, A.L., Francesconi, M. and Garcia-Serrano, 1997, Job tenure: Does history matter?, CEPR Discussion Paper No. 1531.
- [3] Burdett, K. and M. Coles (2003). Equilibrium wage-tenure contracts. *Econometrica* 71 (5), 1377-1404.
- [4] Burdett, K. and D. Mortensen (1998). Equilibrium wage differentials and employer size. *International Economic Review* 39 (2), 257-274.
- [5] Davis, S., and Haltiwanger, J., 1992, Gross Job Creation, Gross Job Destruction and Employment Reallocation, *Quarterly Journal of Economics*, 107, pp. 819-863.
- [6] Farber, H.S., Hall, R. and Pencavel, J., 1993, The Incidence and Costs of Job Loss: 1982-91, *Brookings Papers on Economic Activity. Microeconomics*, Vol. 1, pp. 73-132
- [7] Farber, H.S., 1994, The Analysis of Interfirm Worker Mobility, *Journal of Labor Economics*, 12 (4), pp. 554-593
- [8] Gallant, R. A. and Tauchen, G., 1996, Which Moments to Match?, *Econometric Theory*, 12, pp. 657-681.
- [9] Gibbons, R., Katz, L., Lemieux, T., Parent, D., 2005, Comparative Advantage, Learning, and Sectoral Wage Determination, *Journal of Labor Economics*, 23 (4), pp. 681-723
- [10] Gorry, A., 2009, Labor Market Experience and Worker Flows, Unpublished, University of Chicago
- [11] Gouriou, C., A. Monfort and E. Renault, 1993, Indirect Inference, *Journal of Applied Econometrics*, 8, pp. S85-S118

- [12] Gourieroux, C. and Monfort, A., 1996, *Simulation Based Econometric Methods* (Oxford: Oxford University Press).
- [13] Hall, R.E., 1982, The Importance of Lifetime Jobs in the U.S. Economy, *American Economic Review*, 72 (4), pp. 716-24.
- [14] Jovanovic, B., 1979, Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87 (5), pp. 972-990.
- [15] Lockwood, B., 1991, Information Externalities in the Labour Market and the Duration of Unemployment, 58 (4), pp. 733-753
- [16] Nagypal, Eva, 2007, Learning by Doing vs. Learning About Match Quality: Can We Tell Them Apart?, *Review of Economic Studies*, 74, pp. 537-566
- [17] Papageorgiou, T., 2009, *Learning Your Comparative Advantage*, Unpublished, Penn State University
- [18] Pissarides, C. A. (1994). Search unemployment with on-the-job search. *Review of Economic Studies* 61 (3), 457-75.
- [19] Pries, M.J., 2004, Persistence of Employment Fluctuations: A Model of Recurring Job Loss, *Review of Economic Studies*, 71 (1), pp. 193-215
- [20] Pries, M. and Rogerson, R., 2005, Hiring Policies, Labor Market Institutions, and Labor Market Flows, *Journal of Political Economy*, Vol. 113(4), pp. 811-839.
- [21] Smith, A. A., 1993, Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions, *Journal of Applied Econometrics*, 8, pp. 63 - 84
- [22] Tauchen, G., 1997, New Chi-Squared Methods in Empirical Finance, in D. M. Kreps and K. F. Wallis (eds.) *Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress* (Cambridge: Cambridge University Press) 279-317.
- [23] Topel, R.H. and Ward, M.P., 1992, Job Mobility and the Careers of Young Men, *Quarterly Journal of Economics*, 107(2), pp. 439-79.

2.10 Appendix A

2.10.1 Proof of Proposition 2.1

First we show that the surplus function $S(\cdot)$ is linearly increasing in $\hat{\pi}$. We make use of the definition of $S(\cdot)$ as stated in 2.33. This gives

$$\begin{aligned} \frac{\partial S(k, i, \hat{\pi})}{\partial \hat{\pi}} = & y^g - y^b + \beta \rho_k (1 - \sigma) \sum_{i' \in I} \mu^e(i' | i) \left\{ (1 - \varphi) \frac{\partial S(k', i', \hat{\pi})}{\partial \hat{\pi}} + \varphi S^g(k', i') \right\} \\ & + \beta \rho_k (1 - \sigma) \varphi \left[\sum_{i' \in I} \mu^e(i'^u(k', i'^u(k', i))) \right] \end{aligned} \quad (2.24)$$

$y^g - y^b$ is positive by definition and $S^g(\cdot) > 0$ holds whenever $y^g > b$. Furthermore, $\Gamma^u \geq 0$ for all (i, k) whenever $b \geq 0$, therefore the expression in the lower line is non-negative since $\sum_{i' \in I} \mu^e(i'^u(k', i))$ is a convex combination of $\Gamma^u(k', 1)$ and $\Gamma^u(k', 2)$ and we have that $\Gamma_u(k'^u(k', 1))$. Given this, it remains to find an expression for $\frac{\partial S(k', i', \hat{\pi})}{\partial \hat{\pi}}$. This can be done by recursion. Starting at the terminal date K and using $\Gamma^u(K, i) = b$ we get

$$S(K, i, \hat{\pi}) = y^g \hat{\pi} + y^b (1 - \hat{\pi}) - b \quad (2.25)$$

from which it follows that $\frac{\partial S(K, i, \hat{\pi})}{\partial \hat{\pi}} = y^g - y^b > 0$. Inserting this into (2.24) proves that $\frac{\partial S(K-1, i, \hat{\pi})}{\partial \hat{\pi}} > 0$. By recursion we thus establish that $\frac{\partial S(k, i, \hat{\pi})}{\partial \hat{\pi}}$ is increasing in $\hat{\pi}$ for all k . Next we prove that $S(\cdot)$ is linear in $\hat{\pi}$. Using 2.24 we find

$$\frac{\partial^2 S(k, i, \hat{\pi})}{\partial \hat{\pi}^2} = \beta \rho_k (1 - \sigma) (1 - \varphi) \sum_{i' \in I} \mu^e(i' | i) \frac{\partial^2 S(k', i', \hat{\pi})}{\partial \hat{\pi}^2} \quad (2.26)$$

Applying recursion again and using $\frac{\partial^2 S(K, i, \hat{\pi})}{\partial \hat{\pi}^2} = 0$ establishes the result. To prove that $S(k, i, 0) < 0$, we make use of $J(k, i, 0) = (1 - e)S(k, i, 0)$ and show that $J(k, i, 0) < 0$. We proceed by stating

$$J(k, i, 0) = y^b - w(k, i, 0) + \beta (1 - \sigma) \rho_k (1 - \varphi) \sum_{i' \in \mathcal{I}} \mu_{i|i'}^e J(k', i', 0) \quad (2.27)$$

Notice that $J(K, i, 0) = ey^b - b$ which is negative for $b > y^b$. We substitute $J(K, i, 0) = ey^b - b$ into $J(K-1, i, 0)$ which becomes $J(K-1, i, 0) = y^b - w(K-1, i, 0) + \beta (1 - \sigma) \rho_k (1 - \varphi) (ey^b - b)$. $J(K-1, i, 0) < 0$ follows immediately from $w(K-1, i, 0) > b$. Using this result we can apply recursion to find that $J(k, i, 0) < 0$ for all (k, i) . To prove that $J(k, i, 1) > 0$ we make use of $J(k, i, 1) = J^g(k, i)$, hence it remains to show that $J^g > 0$ for all (k, i) . $J^g(k, i)$ is given by

$$J^g(k, i) = y^g - w(k, i, 1) + \beta(1 - \sigma)\rho_k(1 - \varphi) \sum_{i' \in \mathcal{I}} \mu_{i|i'}^e J^g(k', i') \quad (2.28)$$

Notice that $J^g(K, i) = ey^g - b$ which is positive whenever $ey^g > b$. We substitute $J^g(K, i)$ into $J^g(K - 1, i)$ which becomes $J^g(K - 1, i) = y^g - w(K - 1, i, 1) + \beta(1 - \sigma)\rho_k(1 - \varphi)(ey^g - b)$. $J^g(K - 1, i) > 0$ follows immediately from $w(K - 1, i, 0) < y^g$. Applying recursion to this expression we can establish the result that $J^g(k, i) = J(k, i, 1) > 0$ for all (k, i) .

2.10.2 Proof of Proposition 2.4

Let \mathcal{A}^0 and \mathcal{A}^1 denote two functions that, respectively, represent the expressions on the left-hand side and on the right-hand side of Equation (2.10). Those functions are given by

$$\mathcal{A}^0(\hat{\pi}) = \hat{\pi}S'(\hat{\pi})[1 - \beta\rho(1 - \sigma)(1 - \varphi)] + y^b$$

$$\mathcal{A}^1(\hat{\pi}) = b + \beta\rho\tilde{\lambda}(1 - e)S'(\hat{\pi}) \int_{\hat{\pi} \geq \hat{\pi}} (\hat{\pi} - \hat{\pi})dG^i(\hat{\pi})$$

\mathcal{A}^0 is strictly increasing for all $\hat{\pi} \in [0, 1]$ since

$$\mathcal{A}_{\hat{\pi}}^0 = \frac{[1 - \beta\rho(1 - \sigma)]^2}{1 - \beta\rho(1 - \sigma)(1 - \varphi\hat{\pi})} S'(\hat{\pi}) > 0$$

with intercepts $\mathcal{A}^0(0) = y^b$ and $\mathcal{A}^1 = y^{g34}$. On the other hand \mathcal{A}^1 is strictly decreasing for all $\hat{\pi} \in [0, 1]$ since

$$\mathcal{A}_{\hat{\pi}}^1 = -\beta\rho\tilde{\lambda}(1 - e) \left[\frac{\beta\rho\varphi(1 - \sigma)}{1 - \beta\rho(1 - \sigma)(1 - \varphi\hat{\pi})} S'(\hat{\pi}) + (1 - G^i(S'(\hat{\pi}))) \right] < 0$$

with intercepts $\mathcal{A}^1(0) = b + \beta\rho\tilde{\lambda}(1 - e)\alpha S'(0)$ and $\mathcal{A}^1(1) = b$. Given the restrictions on the parameters and on y^g and y^b , we know that $\beta\rho\tilde{\lambda}(1 - e)\alpha S'(0) > 0$. Moreover if $y^g > b$ and $b + \beta\rho\tilde{\lambda}(1 - e)\alpha S'(0) > y^g$ there exists at least one $\hat{\pi} \in (0, 1)$ for which $\mathcal{A}^0(\hat{\pi}) = \mathcal{A}^1(\hat{\pi})$. Uniqueness follows from the fact that $\mathcal{A}_{\hat{\pi}\hat{\pi}}^0 < 0$ and $\mathcal{A}_{\hat{\pi}\hat{\pi}}^1 > 0$. Notice that restricting y^g , y^b and b so that the condition $y^g > b > y^b$ is fulfilled is actually sufficient for the existence of an unique interior equilibrium. To ensure that it is always optimal to dissolve bad matches the following condition needs to hold in equilibrium

$$y^b \leq b + \beta\rho\tilde{\lambda}(1 - e)S'(\hat{\pi}) \int_{\hat{\pi} \geq \hat{\pi}(i)} (\hat{\pi} - \hat{\pi}(i))dG^i(\hat{\pi})$$

³⁴Notice that $\mathcal{A}_{\hat{\pi}}^0$ denotes the partial derivative of \mathcal{A}^0 with respect to $\hat{\pi}$.

2.11 Appendix B: Threshold signals and threshold beliefs

This section analyzes the mapping between the threshold beliefs $\hat{\pi}$ and the corresponding signal γ . For convenience we call the signal which induces beliefs equal to the threshold value $\hat{\pi}$ as the threshold signal, which we denote by $\hat{\gamma}$. Take two workers of type i and j where $j > i$, hence type i worker emits a more noisy signal $\sigma_{\epsilon,i}^2 > \sigma_{\epsilon,j}^2$. Therefore, $\hat{\pi}(j) > \hat{\pi}(i)$. Recall that $\hat{\pi} = (1 - e_i)\bar{\alpha} + e_i\gamma$. It follows that for $\hat{\pi}(i) \leq \bar{\alpha}$ the relation is $\underline{\gamma}(i) < \underline{\gamma}(j)$. The condition says that if the threshold belief for the type i worker is below or equal the mean value of beliefs, α , then the minimum necessary signal for type i is always lower than for the worker with the more precise signal. In other words, if agents accept forming matches that are based on relatively poor beliefs (reflected by a hiring threshold that is below the mean belief α) then they prefer these particular matches to be created on the basis of a noisy rather than a precise signal. This can be explained by the fact that a match that is based on a bad, yet precise, signal is more likely to turn out to be bad than a match with a bad, yet noisy, signal. The high amount of noise in the signal increases the likelihood that a bad-signal match is in fact good. Therefore, agents accept a lower threshold signal $\underline{\gamma}$ for a match creation.

For the case $\hat{\pi}(i) > \alpha$ matters are less clear-cut. There are two counteracting effects. Whenever $\hat{\pi}(i) > \alpha$, then type j workers require a lower signal, γ , to achieve the same value $\hat{\pi}$ as type i workers. This is implied as the relative advantage of type j workers, in terms of the informativeness of their signal. A good signal that is informative is more likely to indicate a good match than a good signal that is noisy.

On the other hand, we show that type j workers face a higher threshold for their beliefs. To achieve this higher threshold workers naturally require a better signal γ . Which of these two effects dominates depends how far apart the threshold beliefs $\hat{\pi}(i)$ and $\hat{\pi}(j)$ are from each other. If $\hat{\pi}(j)$ is sufficiently close to $\hat{\pi}(i)$, then the first effect dominates and we get that $\underline{\gamma}(j) < \underline{\gamma}(i)$, i.e. the minimum signal required for a successful match creation is lower for type j workers. If, however, $\hat{\pi}(j)$ is far from $\hat{\pi}(i)$ then the reverse is true. Hence, there is a certain distance between the respective threshold values for which the two effects exactly balance, implying that $\underline{\gamma}(i) = \underline{\gamma}(j)$. Let us express the threshold for type i as the deviation from the mean, $\hat{\pi}(i) = \bar{\alpha} + \Delta$ with $\Delta \geq 0$. Using that notation one can show that the critical distance is given by $\frac{\sigma_{\epsilon,i}^2 - \sigma_{\epsilon,j}^2}{\sigma_{\pi}^2}$. Moreover, it is true that

$$\hat{\pi}(j) \begin{cases} < \hat{\pi}^* & \underline{\gamma}(j) < \underline{\gamma}(i) \\ = \hat{\pi}^* & \underline{\gamma}(j) = \underline{\gamma}(i) \\ > \hat{\pi}^* & \underline{\gamma}(j) > \underline{\gamma}(j) \end{cases}$$

where $\hat{\pi}^* = \bar{\alpha} + \left(\frac{\sigma_{\pi}^2 + \sigma_{\epsilon,i}^2}{\sigma_{\pi}^2 + \sigma_{\epsilon,j}^2} \right) \Delta$. It is easy to see that for $\hat{\pi}(i) \leq \bar{\alpha}$ (i.e. $\Delta \leq 0$) there is no $\hat{\pi}(j) > \hat{\pi}(i)$ such that $\hat{\pi}(j) \leq \hat{\pi}^*$. To sum up, if the critical threshold for type i is at or below the mean α then type j can never have an advantage in the threshold signal $\underline{\gamma}$. If $\hat{\pi}(i) > \alpha$, then it depends on the exact location of $\hat{\pi}(j)$ with respect to $\hat{\pi}^*$ whether or not type j enjoys an advantage.

2.12 Appendix C: Computing aggregate variables in the simple model

2.12.1 The type-specific unemployment rate

In each period, a fraction of agents, $(1 - \rho)$, die and exit the model. To keep the size of the population constant we need the same amount of agents entering the market. Labor market entrants are initially unemployed. Remember that the simple version of the model abstracts from the transition of agents' types. For new labor market entrants, we assume that they are split equally between the varioud types. Consequently, a fraction $(1 - \rho)/n_{\mathcal{I}}$ of the total new labour market entrants goes to each of the types. $n_{\mathcal{I}}$ denotes the total number of types in the economy.

Let u^i and r^i denote, respectively, the mass of unemployed type i workers and the mass of employed type i workers that are in a match with known quality. Furthermore, let $n^i(\hat{\pi})$ denote the mass of employed type i workers that are in a match with beliefs $\hat{\pi}$. The total number of new matches for each worker type is $\tilde{\lambda}[1 - G^i(\hat{\pi}^i)]u^i$ and the mass of new matches for each value of beliefs is $\lambda u^i g^i(\hat{\pi})$, where $g^i(\hat{\pi})$ is the type-specific probability density function of the beliefs. The laws of motion for u^i , r^i and $n^i(\hat{\pi})$ are given by

- Law of motion for unrevealed matches for each value of beliefs:

$$n_{t+1}^i(\hat{\pi}) = (1 - \sigma)\rho(1 - \varphi)n_t^i(\hat{\pi}) + \tilde{\lambda}\rho u^i g^i(\hat{\pi}) \quad (2.29)$$

The inflow is given by the the mass of new matches for each signal, given by $\tilde{\lambda}u^i g^i(\hat{\pi})$, whereas the outflow is given by the surviving matches for which the match quality gets revealed.

- Law of motion for revealed matches

$$r_{t+1}^i = r_t^i(1 - \sigma)\rho + \varphi(1 - \sigma)\rho \int_{\hat{\pi}^i} \hat{\pi} n^i(\hat{\pi}) d\hat{\pi} \quad (2.30)$$

The inflow is given by the mass of matches whose quality is revealed to be good.

- Law of motion for unemployed of type i

$$u_{t+1}^i = u_t^i \rho [1 - \tilde{\lambda}(1 - G^i(\hat{\pi}))] + \sigma \left(\int_{\hat{\pi}^i} n^i(\hat{\pi}) d\hat{\pi} + r_t^i \right) \rho + \varphi(1 - \sigma) \rho \int_{\hat{\pi}^i} (1 - \hat{\pi}) n^i(\hat{\pi}) d\hat{\pi} + (1 - \rho) I^i \quad (2.31)$$

The flow out of unemployment, represented by $\tilde{\lambda}(1 - G^i(\hat{\pi}))$, consists of workers that encounter a vacancy and generate a signal that is above the respective reservation signal. The inflow is made up by matches that are destroyed either exogenously or by the revelation of mismatch.

In a stationary equilibrium $x_{t+1}^i = x_t^i = x^i$ for $x \in \{r, n, u\}$, hence we express $n^i(\hat{\pi})$ and r^i as

$$n^i(\hat{\pi}) = \tilde{a} u^i g^i(\hat{\pi}) \quad r^i = \tilde{c} \int_{\hat{\pi}^i} \hat{\pi} n^i(\hat{\pi}) d\hat{\pi} \quad (2.32)$$

where $\tilde{a} = \frac{\tilde{\lambda}\rho}{1 - (1 - \sigma)\rho(1 - \varphi)}$ and $\tilde{c} = \frac{\varphi(1 - \sigma)\rho}{1 - \rho(1 - \sigma)}$. Combining Equations (2.32) and (2.31) yields expression (2.11) in the text.

2.12.2 Match destruction

The rate of match destruction, Φ^i , is computed as the ratio of bad matches to the total number of matches with unknown quality, i.e. $\Phi^i = \frac{\int_{\hat{\pi}^i} (1 - \hat{\pi}) n^i(\hat{\pi}) d\hat{\pi}}{\int_{\hat{\pi}^i} n^i(\hat{\pi}) d\hat{\pi}}$. To arrive at the expression used in the text, we make use of the steady state value of $n^i(\hat{\pi}) = \tilde{a} u^i g^i(\hat{\pi})$.

2.13 Appendix D

2.13.1 The joint surplus function

The definition of the joint surplus of a match with $(k, i, \hat{\pi})$ is given by $S(k, i, \hat{\pi}) = J(k, i, \hat{\pi}) + \Gamma^e(k, i, \hat{\pi}) - \sum_{i'} \mu_{i|i'}^b \Gamma^u(k, i')$. Using the value functions (2.5) and (2.6) to substitute for $J(k, i, \hat{\pi})$ and $\Gamma^e(k, i, \hat{\pi})$ and the surplus sharing rule $\epsilon \left(\Gamma^e(k, i, \hat{\pi}) - \sum_{i'} \mu_{i|i'}^b \Gamma^u(k, i') \right) = (1 - \epsilon)J(k, i, \hat{\pi})$ we arrive at the following expression for S

$$\begin{aligned} S(k, i, \hat{\pi}) = & E(y|\hat{\pi}) \\ & + \beta \rho_k (1 - \sigma) \sum_{i' \in I} \mu^e(i'|i) \{ (1 - \varphi) S(k', i', \hat{\pi}) + \varphi \hat{\pi} S^g(k', i') \} \\ & + \beta \rho_k (1 - \sigma) [1 - \varphi(1 - \hat{\pi})] \sum_{i' \in I} \mu^e(i'|i) \sum_{i'' \in I} \mu^b(i''|i') \Gamma^u(k', i'') \\ & + \beta \rho_k [\varphi(1 - \sigma)(1 - \hat{\pi}) + \sigma] \sum_{i' \in I} \mu^b(i'^u(k', i')) \\ & - \sum_{i' \in I} \mu^b(i'^u(k, i')) \end{aligned} \quad (2.33)$$

2.13.2 Wage formation

Wages are determined by bilateral Nash bargaining between the worker and the firm. In particular, the wage is chosen to maximize the Nash product, hence it has to solve

$$w = \arg \max J(k, i, \hat{\pi})^e \left(\Gamma^e(k, i, \hat{\pi}) - \sum_{i'} \mu_{i|i'}^b \Gamma^u(k, i') \right)^{1-e} \quad (2.34)$$

where e is the fraction of the total surplus that goes to the firm. The optimality condition associated with (2.34) is given by

$$e \left(\Gamma^e(k, i, \hat{\pi}) - \sum_{i'} \mu_{i|i'}^b \Gamma^u(k, i') \right) = (1 - e) J(k, i, \hat{\pi}) \quad (2.35)$$

Using (a) the value functions (2.5) and (2.6) to substitute for $J(k, i, \hat{\pi})$ and $\Gamma^e(k, i, \hat{\pi})$ and (b) the definition of the surplus function (2.33) we arrive at the following expression for the wage function

$$\begin{aligned} w(k, i, \hat{\pi}) = & E(y|\hat{\pi}) \\ & - e \left[S(k, i, \hat{\pi}) - \beta \rho_k (1 - \sigma) \sum_{i' \in I} \mu_{i|i'}^e \{ (1 - \varphi) S(k', i', \hat{\pi}) + \varphi \hat{\pi} S^g(k', i') \} \right] \end{aligned} \quad (2.36)$$

2.14 Appendix E

Actual hazard rate, by job tenure and by age at match formation						
Age / Tenure	[0-1)	[1-2)	[2-3)	[3-5)	[5-7)	[7-10)
[18-20)	1.1282	0.6932	0.4868	0.3513	0.2336	0.1841
[20-23)	1.0275	0.6328	0.4431	0.3139	0.2117	0.1748
[23-26)	0.9514	0.5793	0.3990	0.2940	0.1910	0.1512
[26-30)	0.8376	0.5032	0.3563	0.2556	0.1656	0.1314
[30-34)	0.6963	0.4186	0.2915	0.2065	0.1433	0.1095
[34-38)	0.5846	0.3593	0.2459	0.1773	0.1129	0.0936
[38-	0.5362	0.3340	0.2267	0.1627	0.1054	0.0874

Estimates of workers' actual hazard rate of separating from a job for selected age and tenure groups. Source: Own calculations based on data from the *NLSY* 79.

Table 2.13

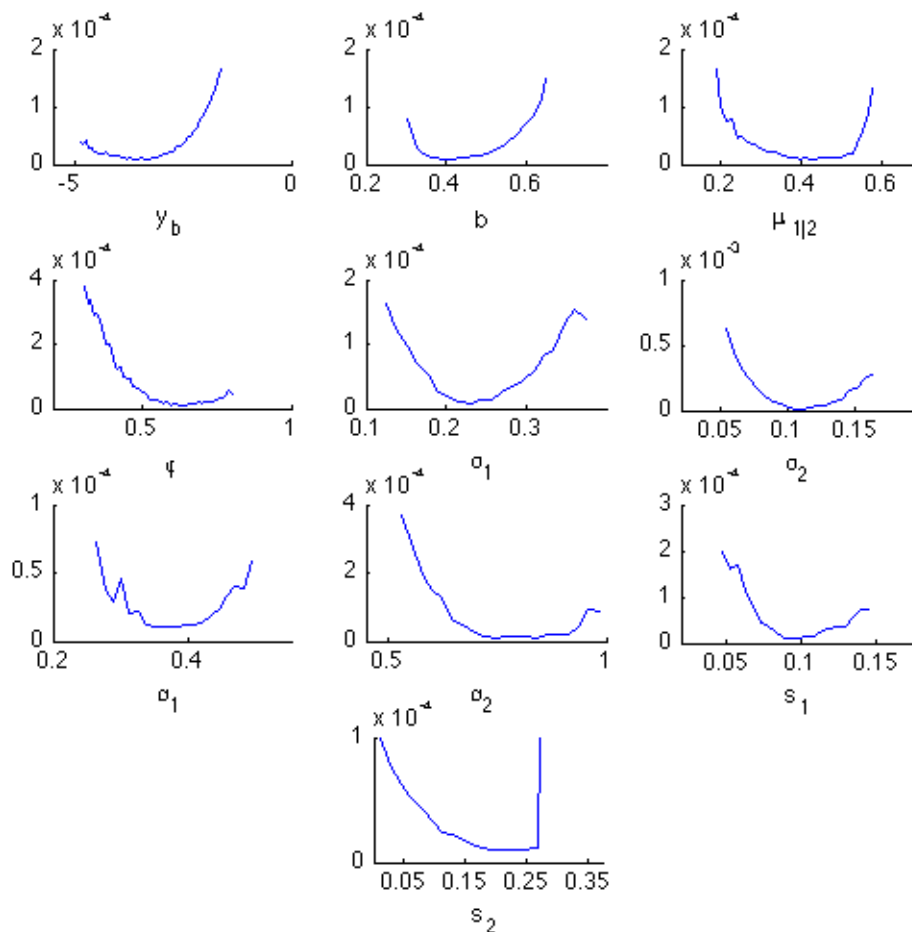
Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.49 (0.57)	0.67 (0.63)	0.71 (0.72)	0.84 (0.79)
2 years	0.31 (0.43)	0.51 (0.51)	0.57 (0.60)	0.58 (0.67)
3 years	0.21 (0.37)	0.42 (0.43)	0.43 (0.52)	0.56 (0.59)
4 years	0.15 (0.29)	0.33 (0.37)	0.39 (0.45)	0.47 (0.52)
5 years	0.13 (0.25)	0.29 (0.33)	0.35 (0.41)	0.44 (0.46)

Job retention rates for workers with one year of tenure in the current job: Rows represent different retention periods, columns indicate workers' age at the time of recruitment. Numbers in bold (normal font) represent the model (data).

Table 2.14

2.15 Appendix F: Accuracy of the estimation results

Figure 2.8



The value of the objective function (stated in Equation (2.20)) evaluated at the optimal estimator for all but one parameter. The remaining parameter is varied over a range of feasible values.

Chapter 3

On the Determinants, Nature and Consequences of Job Separation

Abstract

This paper uncovers and explores the main elements related to the causes and consequences of job separation. In particular, it empirically addresses the following three questions using data on white males from the National Longitudinal Survey of Youth (1979): (a) What are the determining factors of separation hazard of employment relationships? (b) Which employer-employee matches are more likely to dissolve due to a layoff rather than a voluntary separation? (c) What are the effects of a voluntary separation and a layoff on re-employment wages? This paper offers a set of answers to those questions, some of which confirm previous findings in the literature, whereas others provide new facts and insights and also pose a challenge to conventional theoretical labor market models, such as the standard job ladder model. The controversial elements will be used as a point of departure for future research on individuals' labor market dynamics.

JEL Classification: C33, C41, E24, J63

Keywords: Job separation, Reemployment, Quit gains, Firing losses, NLSY 79

3.1 Introduction

This paper addresses three sets of questions related to the determinants and the consequences of job separation:

Question 1: What are the determining factors of separation hazard of employment relationships? In the data we observe a tremendous amount of labor reallocation. All in all, 6.7% of all U.S. employee-employer matches are dissolved in an average month. At the same time, 4.3% of the civilian noninstitutional population aged 16 and over began to work in a new job (Fallick and Fleischman (2004)). The duration distribution of employment spells in the US labor market is, therefore, highly skewed to the left but despite the high turnover it is clearly non-degenerate. Many of the newly created matches survive only for a short period of time, a significant proportion, however, are more durable and turn into medium and long term jobs. This raises an obvious question, namely: what are the underlying factors and determinants that make some of the new matches likely to survive longer than others. Job survival is not a purely random process and we may presume that it is in fact governed by a given (possibly very large) set of determinants.

One objective of this paper is to identify worker- and match-specific factors that are causal for the observed systematic differences in the actual duration of employment spells. To achieve this goal, I apply econometric methods that are tailored to the analysis of duration problems. In particular, I estimate a discrete-time hazard rate model that is augmented by a set of possible explanatory variables. The variables considered in the analysis mainly include factors that relate to the characteristics of (a) the employed worker (e.g. age, education, etc.), (b) the worker's current employment spell (e.g. the wage) and (c) the worker's recent labor market history (reason of separation of last job, number of weeks not employed, etc.). The second question the paper addresses is a follow-up on the first one:

Question 2: Which employer-employee matches are more likely to dissolve due to a layoff rather than a voluntary quit? From a worker's viewpoint, the separation from the current employer can be voluntary or involuntary. I will refer to the first case as *quit* and to the latter as *layoff*. The nature of a separation is crucial for the economic well-being of the individual involved. A layoff is typically perceived as an adverse (and often unexpected) economic shock against which markets provide only incomplete insurance. Moreover, it is associated with serious consequences that may come in the form of income and skill loss, especially when it leads to prolonged periods of unemployment. On the other

hand, a quit is considered as something positive, as it often leads to an immediate (and foreseen) transition to another job and thereby to an improvement in an individual's economic situation. The second goal of the paper is to identify factors which help to assess whether a given match is likely to dissolve due to a layoff rather than a quit. This issue is analyzed by using an estimated binary response model which relates the probability of observing an involuntary separation to a pre-defined set of explanatory variables. The third question of the paper closes the circle:

Question 3: What are the effects of a voluntary quit and a layoff, respectively, on re-employment wages? After a job separation, workers eventually become reemployed. In the case of a quit, reemployment is often equivalent to a job-to-job transition. Since those transitions are mostly initiated by the worker, we would generally expect voluntary separations to be followed by wage gains. The impact of a layoff is less clear. It may be associated with the loss of firm- or industry-specific skills, which would generally lead to a lower reemployment wage. On the other hand, a separation can also reflect the break-up of a bad match and the subsequent reallocation of a worker's labor input towards a possibly more productive job. In that case, the wage earned in the new occupation is likely to be higher than the previous. The last part of the paper addresses very briefly the issue of how wages after reemployment compare depending on the nature of the preceding separation.

This paper is purely empirical and as such, conclusions derived herein follow exclusively from an econometric rather than theoretical or structural study. The questions above are addressed using data on U.S. white males from the National Longitudinal Survey of Youth 1979, (henceforth *NLSY 79*). What is the purpose of asking and addressing the aforementioned questions? This study is meant to be a preparatory study which should uncover and explore the main elements related to the causes and consequences of labor market transitions and in particular, job separations and reemployment. It is not intended to be a "stand-alone" work but rather a point of departure for future, mostly theoretical or structural, research. The aim is to gather insights and to establish a set of facts on the process surrounding labor market transitions. The insights and facts that are established by this paper to be, empirically, both important and relevant will be pursued in future research.

For example, the results of this paper pose a severe challenge to the logic and predictions of the standard job ladder model, as used by Pissarides (1994), Burdett and Mortensen (1998) and Burdett and Coles (2003) among others. This challenge stems from a set of em-

empirical regularities detected in the data that are contradictory to the predictions of the job ladder model. (The specific details will be dealt with later). The idea is then to make use of the insights gathered here to address the issue theoretically and in particular, to come up with a modeling device whose predictions are in line with the empirically observed patterns.

The remainder of the paper is structured as follows. Section 3.2 discusses the empirical framework that is used to address the aforementioned questions and issues. Section 3.3 describes the data used in the estimation. This section also provides a detailed explanation of the way in which each of the considered variables is constructed¹. Section 3.4 reports the results obtained from the empirical analysis and Section 3.5 discusses the results and concludes.

3.2 The Empirical Framework

In this section, I discuss the empirical framework that is used to analyze the first two questions formulated at the outset. The third question is not addressed within a formal econometric framework. The reason being that preliminary analysis has revealed that reemployment wages are highly dispersed, even within narrowly defined groups. A formal econometric analysis would be doomed to fail as the expected large standard errors would make the estimates meaningless. Nonetheless, I intend to make a statement about Question 3 and to that end, I report certain descriptive statistics, which are meant to provide a rough illustration of the issue.

3.2.1 Addressing Question 1: The determinants of job separation

The aim is to detect factors that are causal for the observed differences in job duration. To that end, I estimate a discrete-time hazard rate model that formalizes the mapping from observable factors (such as worker characteristics) to the job separation probability. The specification used here is similar to that in Duernecker (2009) where the model serves as the auxiliary model in the structural estimation. The main difference is the set of explanatory variables considered in the estimation.

The data that is used in the estimation consists of observations of individuals' employment spells. Each observation is indexed by $j \in \{1, 2, \dots, J\}$, where J denotes the total number of observations. For the estimation I divide the time line into N intervals that are given by $[0, \tau_1)$, $[\tau_1, \tau_2)$, ..., $[\tau_{N-1}, \tau_N)$. Intervals are indexed by the subscript $n \in \{1, 2, \dots, N\}$. Let

¹The raw and processed data that is used in this study is available upon request.

t_j denote the duration of employment spell j . The hazard rate of separating from job j in interval n is modeled as a piecewise constant function given by

$$\lambda(t; \mathbf{x}_j, \beta) = \kappa(\mathbf{x}_j, \eta) \lambda_n, \quad t \in [\tau_{n-1}, \tau_n) \quad (3.1)$$

β is a vector of unknown parameters and $\kappa(\mathbf{x}_j, \eta) > 0$ is a function of observable worker characteristics. Those characteristics are captured by the vector \mathbf{x}_j , whereas η is a vector of unknown parameters. The baseline hazard for each interval is given by λ_n . This specification implies a constant hazard within a particular interval n , yet it allows the hazard to differ across intervals. Using the hazard rate defined in (3.1) one can construct the log likelihood function of observation j .

$$\begin{aligned} l(n_j, \mathbf{x}_j, c_j; \beta) = & - \sum_{h=1}^{n_j-1} \kappa(\mathbf{x}_j, \eta) \lambda_h (\tau_h - \tau_{h-1}) \\ & + c_j \log \{1 - \exp [-\kappa(\mathbf{x}_j, \eta) \lambda_{n_j} (\tau_{n_j} - \tau_{n_j-1})]\} \end{aligned} \quad (3.2)$$

The first part is the probability the job lasts until n_j and the second part is the probability that it ends in the interval n_j . Notice that the latter part is non-zero only when the observation is uncensored ($c_j = 1$) and therefore represents a true job loss. The log likelihood function for the entire sample is obtained by summing over all $j = 1, \dots, J$ employment spells. It is given by

$$\mathcal{L}(\mathbf{y}_J; \beta) = \sum_{j=1}^J l(n_j, \mathbf{x}_j, c_j; \beta) \quad (3.3)$$

where $\mathbf{y}_J = \{n_j, \mathbf{x}_j, c_j\}_{j=1}^J$. Using the information on (n_j, \mathbf{x}_j, c_j) that is available for each employment spell j we can estimate the vector of unknown parameters β by maximizing the log likelihood function in (3.3).

$$\hat{\beta}_J = \arg \max_{\beta} \mathcal{L}(\mathbf{y}_J; \beta) \quad (3.4)$$

Before estimating β one has to take a stand on the functional form of $\kappa(\mathbf{x}_j, \eta)$ and which explanatory variables to include in \mathbf{x} . The functional form for $\kappa(\cdot, \cdot)$ is chosen to be $\kappa(\mathbf{x}_j, \eta) = e^{\eta \mathbf{x}_j}$ and the vector of explanatory variables, considered in the estimation, consists of three sets of categories.

1. Match-specific characteristics

Variables falling into this category are meant to describe the current state of the employer-employee match. The state is likely to influence the separation decision. Therefore, any variable that characterizes the match along a particular dimension

will contain information about the separation hazard. Here I consider $\Delta\omega_{j,in}$: The percentage deviation of the **initial** wage payed in spell j from the average initial wage payed to individuals that are similar in characteristics to the worker in spell j .

2. Worker-specific characteristics

age: The worker's age at the point in time when the match was formed. Duernecker (2009) shows that a worker's experience in the labor market, as proxied by his age, is an important determinant of job duration. *Coll*.: A dummy variable that is unity if the individual holds a college degree and zero otherwise.

3. Characteristics that relate to the workers (recent) labor market history

$\Delta\omega_{in}^{prev}$: The percentage deviation of the **initial** wage payed in job j from the last wage payed in the worker's **previous** occupation. $\Delta\omega_{in}^{prev}$ being positive indicates that the worker made a leap upwards in the wage ladder. \bar{D} : A dummy variable that is unity (zero) if the separation from the previous job was due to a layoff (quit). τ_{prev} : The worker's total tenure on the previous job. \tilde{w} : The duration of non-employment. This variable effectively counts the number of weeks between separation from the previous job and start of the current one.

3.2.2 Addressing Question 2: Investigating the nature of job separations

For this part of the analysis, I restrict the attention to those employment spells for which we observe a true separation. Whenever a respondent in *NLSY* reports that he stopped working for a particular job he is also asked if this was due a quit or a layoff². We can therefore assign to each employment spell j a binary variable, denoted by z_j , that is unity if for job j we observe a layoff and zero otherwise. The goal is then to find the effect of a set of variables on the probability of observing a layoff. The variables I consider here, again, consist of worker- and match-specific characteristics and broadly correspond to the three categories mentioned above.

More formally: The task is to estimate the vector of parameters, captured by the $K \times 1$ vector θ , that defines

$$P(z_j = 1|\bar{\mathbf{x}}_j) = G(\bar{\mathbf{x}}_j\theta) \quad (3.5)$$

$\bar{\mathbf{x}}$ denotes the $J \times K$ matrix of explanatory variables. I assume that G is the cumulative distribution function of a standard normal distribution which makes the model in (3.5) a

²More precisely, the respondent is asked to classify the nature of the separation using a list of predefined categories. I will be more specific about those categories in Section 3.3.

Probit model. θ can be estimated by conditional Maximum Likelihood. The density of z_j given $\bar{\mathbf{x}}_j$ can be written as

$$f(z|\bar{\mathbf{x}}_j; \theta) = [G(\bar{\mathbf{x}}_j\theta)]^z [1 - G(\bar{\mathbf{x}}_j\theta)]^{1-z} \quad z = 0, 1 \quad (3.6)$$

From (3.6) we can derive the log-likelihood of observation j as function of the parameters θ and the observed data $(\bar{\mathbf{x}}_j, z_j)$

$$l_i(z_j, \bar{\mathbf{x}}_j; \theta) = z_j \log [G(\bar{\mathbf{x}}_j\theta)] + (1 - z_j) \log [1 - G(\bar{\mathbf{x}}_j\theta)] \quad (3.7)$$

Let's define $\mathbf{y}_J = \{z_j, \bar{\mathbf{x}}_j\}_{j=1}^J$. The log likelihood of the sample is therefore given by $\mathcal{L}(\mathbf{y}_J; \theta) = \sum_{j=1}^J l_j(z_j, \bar{\mathbf{x}}_j; \theta)$ and the probit estimator of θ , denoted $\hat{\theta}$, maximizes this log likelihood. More formally

$$\hat{\theta} = \arg \max_{\theta} \mathcal{L}(\mathbf{y}_J; \theta) \quad (3.8)$$

Under the standard regularity conditions, $\hat{\theta}$ is consistent and asymptotically normal. The covariates included in $\bar{\mathbf{x}}$ correspond to those in \mathbf{x} used in the previous section. In addition, I consider two more variables, namely: (1) τ : The workers total tenure on the current job by the time of the separation and (2) $\Delta\omega_{in}^{last}$: The percentage deviation of the **last** wage paid in job j from the **initial** wage paid in the same occupation. This variable indicates the total amount of wage growth a worker has experienced in job j

3.3 The Data

The data used in the estimation comes from the National Longitudinal Survey of Youth 1979 (*NLSY* 79). The *NLSY* 79 is a representative sample of 12,686 young men and women who were 14 to 22 years of age when first surveyed in 1979. The cohort was interviewed annually through 1994. Since 1994, the survey has been administered biennially. The last survey that is available was published in 2006. The entire data set used in the estimation thus consists of 27 individual surveys. The *NLSY* 79 contains demographic variables, labor market data and information on individuals' wealth and consumption.

The data set I construct consists of employment spells, each denoted by j , and a set of individuals characteristics associated with the respective employment spell. To guarantee a reasonable degree of homogeneity across individuals I exclude certain observations from the data set. In particular, the final data set consists of observations of white males that

had their first job at the age of 18 or later. Moreover, I exclude part time jobs and consider just those employment spells for which individuals worked for ≥ 30 hours per week. Lastly, I do not consider employment spells that began while the worker was simultaneously employed also in one or more other jobs. For a spell to be considered, there has to be a reported separation from any previous job. I do this exclusion mainly for homogeneity reasons.

The construction of the data set requires finding, for each j , (a) the total length of spell j , denoted t_j , (b) the value of the binary variable c_j that indicates a censored observation, (c) the nature of the separation z_j (only for the Probit estimation) and (d) the values of the relevant covariates \mathbf{x}_j and $\bar{\mathbf{x}}_j$. The interested reader is referred to Duernecker (2009), which describes in detail how the values for t , c *age* and $\Delta\omega_{in}$ are recovered. Here I proceed by explaining how the remaining variables are constructed.

$Coll.$ denotes a binary variable that is unity if the individual employed in job j has completed college and is zero otherwise. This variable is relatively straightforward to construct as each respondent in the *NLSY* is asked about his formal education during every interview.

$\Delta\omega_{in}^{prev}$ is the percentage deviation of the **initial** wage paid in job j from the last wage paid in the worker's **previous** occupation. It is essentially the wage gain/loss the worker makes by moving to a new job. To compute $\Delta\omega_{in}^{prev}$ it is first necessary to identify the job that is preceding the current spell j , secondly to determine the date of separation, third to find the last wage paid and lastly, to compute the percentage deviation to the initial wage in spell j^3 . For each employment spell, I observe the week number of its start and, if applicable, also of the separation. This is the key information needed to identify the job that was preceding job j . If this is accomplished, all that is left is to read off the last wage that was paid in that job and to compute the percentage deviation to the initial wage in spell j .

\bar{D} is a binary variable that is unity if the separation from the previous job was due to a layoff and it is zero if the separation was a quit. For constructing \bar{D} it is useful that in the previous step we have, for each j , already identified the preceding job. Next, we simply record the nature of the separation from that job. In the same way I construct τ_{prev} , which

³Notice that all wages are measured in real terms. Therefore, $\Delta\omega_{in}^{prev}$ is not affected by the duration of non-employment between two consecutive jobs. If wages were measured in nominal terms then any non-zero aggregate (nominal) wage growth would cause the wage gap between two jobs to rise proportionally with the spell of non-employment. This effect is netted out by using real wages. To obtain the real wage, the nominal one is deflated with the weekly CPI deflator.

denotes the worker's total tenure on the previous job. \tilde{w} is the duration of non-employment (measured in weeks) between the separation from the previous job and the start of spell j . To compute \tilde{w} I use the week number of the start of spell j and the week number of the previous separation. \tilde{w} is then simply the difference between the two figures.

z is the dependent variable in the Probit estimation. It is a binary variable that is unity (zero) if employment spell j ends due to a layoff (quit). To classify the nature of a separation, each respondent in the *NLSY* is given a range of possible reasons. For most of the survey years those typically include the following: "Layoff", "Plant closed", "End of temporary or seasonal job", "Discharged or fired", "Quit for pregnancy or family reasons", "Quit to look for another job", "Quit to take another job" and "Quit for other reasons". In the coding of z and of \bar{D} , I consider the first four categories as indicating a layoff, and the latter four as describing a quit. $\Delta\omega_{in}^{last}$ denotes the difference (in %) between the **initial** and the **last** wage payed in job j . Here, I once again make use of the knowledge about the week number of each event. In that way I can identify the exact date of the separation and the last wage that was payed in job j . Computing $\Delta\omega_{in}^{last}$ is then straightforward.

The data set used to estimate the hazard rate model consists of 17,966 observations of employment spells, 87.68% of which are uncensored. The data set used in the Probit estimation consists of 14,835 observations. Censoring is not an issue here, as all the spells for which we do not observe a separation are excluded. Table 3.1 reports the summary statistics, in terms of the mean and the standard deviation, for all the variables included in the estimation. Panel (a) depicts the sample used in estimation of the hazard rate model, whereas Panel (b) represents the data for the Probit regression.

Not surprisingly the mean job duration τ , depicted in the first column, is lower in Panel (b) as all the ongoing jobs are excluded there. The standard deviation of τ is fairly large in both samples, indicating that the pool of jobs consists also of a sizable fraction of medium and long term jobs. The mean of $\Delta\omega_{in}$ is zero. This follows by construction as $\Delta\omega_{in}$ represents the deviation of a worker's initial wage from the average initial wage of the group the worker belongs to. The characteristics defining a group are a worker's age (at match formation) and education. Notice that the dispersion of $\Delta\omega_{in}$ across individuals is fairly large. For both samples, the reported standard deviation is more than 30%.

In both samples, the proportion of jobs ending with a layoff is around 26%, indicating that the majority of separations are voluntary (from the worker's point of view). The next two columns (5 and 6) report the average number of weeks a worker stays non-employed prior to the start of a new employment spell. The columns are distinguished from each other by the nature of the separation from the previous job. Columns 5 and 6 respectively refer to spells of non-employment after a layoff and a quit. In both Panels (a) and (b), the average duration until reemployment is substantially shorter after a quit than a layoff. This is plausible as many of the observed quits are the result of a job-to-job transition. Typically, in those cases the worker already has a job lined up before the separation is effectively implemented. Thus the average period of non-employment is relatively short. Instead, a dismissal often constitutes an unforeseen separation from the existing job. This forces workers to search for a new job, which naturally takes time and so extends the period of non-employment.

	τ (yrs.)	age (yrs.)	$\Delta\omega_{in.}$ (%)	\bar{D} (%)	$\tilde{w} _{fir.}$ (wks.)	$\tilde{w} _{quit}$ (wks.)	$\Delta\omega_{in}^{prev}$ (%)	$Coll.$ (%)	τ_{prev} (yrs.)	$\Delta\omega_{in}^{last}$ (%)
Panel (a): Hazard rate model										
<i>Mean(x)</i>	2.01	26.7		25.82	11.8	6.5	13.2	25.5	1.32	
<i>Std(x)</i>	3.62	6.7	35.37		12.9	11.1	42.9		2.38	
Panel (b): Probit model										
<i>Mean(x)</i>	1.28	25.9		26.66	11.6	6.8	12.1	22.9	1.09	1.8
<i>Std(x)</i>	2.09	6.2	31.41		12.8	11.4	38.8		1.89	21.1
Mean and standard deviation of all variables used in the estimation of the hazard rate model and the Probit model outlined in Section 3.2. Sample consists of U.S. white male workers aged 18-48 years. Data: <i>NLSY 79</i> .										

Table 3.1 Summary Statistics

Column 7 shows that the average wage differential of two consecutive jobs amounts to approximately 13% (Panel (b): $\sim 12\%$). The gain seems remarkable but this number is not very robust as the standard deviations are very large and in this case, they are around 3 times the mean value. Roughly one in four observations involves a worker holding a college degree. This number is just slightly lower in the second data set. The penultimate column shows that the average length of the job that is preceding spell j is 1.32 years in Panel (a)

and 1.09 years in Panel (b). Duernecker (2009) shows that the average job duration increases as workers accumulate experience in the labor market. This is also given here, as can be seen by comparing the first column with the eighth column which depict, respectively, the duration of the current and the preceding job. The most recent job is, on average, more durable than the previous one.

The last column of Table 3.1 reports that the average total wage growth on a given job is 1.8%. However, the average across all employment spells is not very meaningful as some jobs last much longer than others and therefore, the total wage change will naturally differ. The large standard deviation is largely due to that reason. To achieve a better comparison, I normalize the total wage growth by job duration. More specifically, I compute the average monthly wage growth for each employment spell and find that on-the-job wage growth is, on average, 0.42% per month. The standard deviation associated with this figure is fairly large, 5.44%.

3.4 Estimation Results

3.4.1 Results for Question 1: The determinants of job separation

This section reports the results obtained from estimating the hazard rate model in (3.1) - (3.3). Before the model is estimated, one has to clarify some issues regarding the exact specification:

- The first issue concerns fixing the width of the intervals within which the baseline hazards, λ_n , are assumed to be constant. This amounts to choosing N grid points τ_n which define $[0, \tau_1)$, $[\tau_1, \tau_2)$, ..., $[\tau_{N-1}, \tau_N)$. Most separations occur early in an employment relationship. To account for this, I make the intervals finer for short durations and wider for long durations.

The final grid consists of 10 intervals which are separated by the endpoints:

$\{0.25, 0.5, 0.75, 1, 1.5, 2, 3, 5, 7, 10\}$. Each of these points denotes job duration as measured in years. All separations occurring after 10 years or more are, therefore, right-censored.

- I do not treat *age* as a continuous variable but instead, I define a set of different age groups and work with dummy variables. More specifically, I define 7 age groups with endpoints $\{20, 22, 25, 28, 31, 34, 38\}$, the lowest starting at age 18 years. The reference group is that with age > 38 years.

- Preliminary results have shown that there is a significant non-linearity in the effect of weeks of non-employment \tilde{w} on the subsequent separation hazard. To account for this,
- intervals and dummy variables are also used in this case. In particular, I define the following two intervals for \tilde{w} , namely $[0, 2]$, $(2, 10]$ and use the group with $\tilde{w} > 10$ weeks as the reference group.
- Table 3.1 has shown that the period of non-employment differs significantly in length depending on the nature of the preceding separation. Thus, I include an interaction term in the estimation namely, $\bar{D} * \tilde{w}$, which is designed to pick up any effect associated with the previous separation. I also include the interaction term $\bar{D} * \tau_{prev}$ ⁴.

The total number of parameters that are estimated is 28. Those are: (a) the ten baseline hazard rates λ_n and the coefficients associated with (b) seven age dummies age_k , (c) the deviation of the worker's initial wage from the group average $\Delta\omega_{in.}$, (d) the dummy variables indicating the spell of non-employment $\tilde{w}_{0|2}$, $\tilde{w}_{3|10}$, $\bar{D} * \tilde{w}_{0|2}$, $\bar{D} * \tilde{w}_{3|10}$, (e) the wage gap between the current and the last job $\Delta\omega_{in}^{prev}$, (f) the binary variable indicating the worker's education $Coll.$ and (g) the worker's tenure on the previous job τ_{prev} , τ_{prev}^2 , $\bar{D} * \tau_{prev}$, $\bar{D} * \tau_{prev}^2$. Notice that I also consider a quadratic term for τ_{prev} and the interaction term $\bar{D} * \tau_{prev}$. The vector of parameters is estimated by solving the maximization problem stated in (3.4). The results are reported in Table 3.2.

Let us first focus on the estimates of the baseline hazard rates, λ_n . These are all quite accurately estimated and statistically significant at the 99% level of confidence. Comparing across intervals, the estimated values are generally in line with expectations. Hazard rates at short durations are generally very high but decreasing in job tenure. This pattern is intuitive as most of the learning concerning the match quality takes place within the first few periods of an employment relationship. Matches that are revealed to be good continue

⁴Conventional models of the labor market, such as the job ladder model, predict a very different hazard for a worker's current spell depending on the length of the previous job **and** the nature of the separation. In such a model, a quit that is ending a long term job indicates a transition to a new job that must have a higher productivity than the previous one and is, thus, expected to last a relatively long time. However, a dismissal from the same job would mean that the worker falls off the ladder and subsequently faces a period in which he holds a series of short-lived jobs before finding a stable match. On the other hand, a separation from a short-term job implies that the worker is still in the lower part of the job/wage ladder. In that case, the difference in the expected duration of the next job is fairly similar - irrespective of whether the preceding separation is a quit or a layoff.

to exist whereas bad matches break up. Therefore, any job that survives to the medium- and long-run is less likely to be of bad quality, which explains the non-linear decline in the baseline hazard rate.

Coeff.	Value	A.H.	Coeff.	Value	$\Delta_{av.}$	$\Delta_{ref.}$	Coeff.	Value	$\Delta_{av.}$	$\Delta_{ref.}$
$\lambda_{0.25}$	1.038 (0.051)	1.44	age_{20}	0.584 (0.041)	-24.1	-51.1	$\tilde{w}_{0 2}$	-0.223 (0.037)	17.1	33.8
$\lambda_{0.5}$	1.019 (0.051)	1.27	age_{22}	0.557 (0.038)	-21.1	-49.1	$\tilde{w}_{3 10}$	-0.034 (0.041)	-8.3	4.9
$\lambda_{0.75}$	0.769 (0.041)	0.91	age_{25}	0.526 (0.036)	-17.6	-46.8	$\bar{D}\tilde{w}_{0 2}$	0.369 (0.035)	-29.2	-19.1
λ_1	0.648 (0.036)	0.72	age_{28}	0.417 (0.037)	-4.1	-38.1	$\bar{D}\tilde{w}_{3 10}$	0.043 (0.038)	-13.6	-1.37
$\lambda_{1.5}$	0.559 (0.029)	0.58	age_{31}	0.329 (0.039)	7.6	-30.6	$\Delta\omega_{in}^{prev}$	-0.189 (0.024)	2.5 (10%)	12.9 (50%)
λ_2	0.438 (0.025)	0.42	age_{34}	0.204 (0.042)	25.1	-19.3	$Coll.$	-0.261 (0.019)	26.9	38.9
λ_3	0.361 (0.021)	0.34	age_{38}	0.059 (0.042)	46.4	-5.6	τ_{prev}	-0.156 (0.009)	15.9 (1 yr)	75.9 (5 yrs)
λ_5	0.264 (0.015)	0.23	$\Delta\omega_{in.}$	-0.510 (0.031)	6.8 (10%)	35.8 (50%)	τ_{prev}^2	0.006 (0.001)		
λ_7	0.165 (0.011)	0.13					$\bar{D}\tau_{prev}$	-0.051 (0.01)	24.9 (1 yr)	105.8 (5 yrs)
λ_{10}	0.137 (0.011)	0.11					$\bar{D}\tau_{prev}^2$	0.002 (0.001)		

Obs.: 17,966 $\log L = -35,699.8$, LR Tests: H_0 : only λ_i (p: 0.000)

λ_τ denotes the estimate of the baseline hazard rate for the interval with endpoint τ . The remaining coefficients are labeled in the same way as the respective covariate A.H. denotes the average effect. $\Delta_{av.}$ is the change (in %) in expected job duration (of the average worker) implied by a given change in the respective explanatory variables. $\Delta_{ref.}$ is the same for dummy variables using the reference group as benchmark. Standard errors are in parentheses.

Table 3.2 Estimation Results: Hazard Rate Model

I also report the average hazard rate for each job duration. It is computed as the mean hazard rate of all observed individuals using the estimated coefficients. For instance, the mean hazard in the population of separating from an employer in the tenure interval $[\tau_{n-1}, \tau_n)$ (conditional on still being in the job) is given by $\lambda(n) \sum_{j=1}^J e^{\eta \mathbf{x}_j} / J$. Conditional on surviving one year, a match is half as likely to break up than a new match. After two years the separation probability is already more than 70% lower and after 5 years, shrinks by about 85%.

Let us next turn to the coefficient estimates of the explanatory variables. To facilitate the interpretation of the estimates, I provide two measures of comparison. $\Delta_{av.}$ computes the percentage change in the expected job duration implied by a given change in one of the explanatory variables. Since $\kappa(\mathbf{x}, \eta) = e^{\mathbf{x}\eta}$, it obviously matters for the marginal effect at which point $\kappa(\cdot, \cdot)$ is evaluated. To make things comparable, I evaluate the estimated hazard rate function at the mean value of each of the covariates, i.e. I compute the hazard rate (and the expected job duration) for a hypothetical "average worker". Thus, $\Delta_{av.}$ effectively represents the change in the expected job duration of the average worker. For all dummy covariates, I also compute the percentage change in the expected job duration with respect to the reference group (instead of the average worker). This is denoted by $\Delta_{ref.}$.

The expected job duration is straightforward to compute and is derived using the estimated hazard rate function. For a new job it is given by $\int_0^\infty t f(t) dt$ where $f(t) = \lambda(t)(1 - F(t))$ is the probability density function of job duration. $\lambda(t)$ denotes the hazard rate of job separation at duration t and $F(t)$ is the probability that the job survives until t . In our setting, $F(t) = 1 - e^{-\int_0^t \lambda(s) ds}$ and $\lambda(t) = \bar{\lambda}(t) \kappa(\mathbf{x}, \eta)$, where $\bar{\lambda}(t)$ is the implied baseline hazard at t . Thus, the expected job duration is given by⁵

$$q(\mathbf{x}, \eta) = \kappa(\mathbf{x}, \eta) \int_0^\infty t \bar{\lambda}(t) e^{-\kappa(\mathbf{x}, \eta) \int_0^t \bar{\lambda}(s) ds} dt \quad (3.9)$$

Let us next turn to the estimated effect of age on the separation probability. The coefficients for all, but one, age groups are statistically different from zero at the 99% level of confidence. In our setup this implies that a worker aged 34 years or less faces a hazard rate that, *ceteris paribus*, is different from the hazard rate of the reference group (the group aged 38 years and above). In contrast, the hazard of the group aged 34 – 38 years is statistically indistinguishable from the reference group. An individual aged 38 years and above has an expected job duration of 2.99 years. For the young, aged 20 years and below, it is only half

⁵ $\bar{\lambda}(t)$ is computed by linearly interpolating the estimated values of λ_n .

of that value. The gap gradually closes for higher age groups. The age group with 34 – 38 years faces an expected job duration that is only about 6% lower than that of the reference group. Assuming that individuals enter, on average, at age 20 we can infer that the first 10 years of labor market experience raise the predicted job duration by about 40%

The coefficient on $\Delta\omega_{in.}$, which is the differential of an individual's initial wage from the group average, is highly significant and has a positive sign. This is plausible as the wage, especially the initial wage, can be thought of as reflecting the employer's perception about the quality of a match. An employer that is willing to pay an above average wage ($\Delta\omega_{j,in.} > 0$) is probably rather optimistic about the worker and his productivity. If the beliefs are, on average, correct then we would expect jobs with $\Delta\omega_{j,in.} > 0$ to last longer. Indeed, the estimation reveals that jobs with a 10% higher initial wage are expected to last 7% longer than the "average job"⁶. A wage differential of 50% leads to a rise in job duration by 36%.

The spell of non-employment between two consecutive jobs turns out to be fairly important for the survival probability of the new job. The coefficient on $\tilde{w}_{0|2}$ indicates that having a short spell, of at most 2 weeks, increases the expected duration of the current job by 34% relative to a situation in which the time until reemployment is 10 weeks or more. The longer the duration of non-employment, the weaker the (positive) effect on the hazard. For spells of 3-10 weeks the difference to the reference group (10 weeks and more) is already fairly small, namely about 5%. Not only is the duration of non-employment important but also how the preceding job ended. A worker who quits and becomes reemployed within 2 weeks has a 65% higher expected job duration than a worker who was fired. A short period of non-employment, most likely, represents a job-to-job transition. Thus, the large difference in job duration seems quite plausible.

However, one might argue that the returns to searching are increasing in duration (at least up to a certain point)⁷. Thus, the comparison of the 0-2 weeks horizon for quits with the 0-2 weeks horizon for layoffs is meaningful only to a limited extent. To address this concern, I compute the expected job duration after a job-to-job transition (quit + reemployment after 0-2 weeks) and after a layoff. When we consider a longer search horizon after a layoff then the difference in expected job duration drops from 65.4% (q: 0-2 weeks / f: 0-2 weeks)

⁶The expected job duration for the average worker is 1.93 years.

⁷If the unemployed are given more time to search for a suitable match then they are more likely to find a good and, therefore, longer lasting job.

to 35.6% (q: 0-2 weeks / f: 3-10 weeks) and 25.9% (q: 0-2 weeks / f: 10-25 weeks). For longer horizons, the difference always remains above 20%. Thus, the returns to searching (in terms of job stability) are clearly positive but nevertheless, the difference between a job-to-job and a job-to-nonemployment-to-job transition always remains significant.

The wage gain between two consecutive jobs, represented by $\Delta\omega_{in}^{prev}$, has a positive and statistically significant impact on the current hazard. However, the economic effect is rather small. A wage gain of 10% (relative to the mean) reduces the hazard of separating and increases the expected job duration by about 3%. The positive sign is plausible as a wage gain represents an upward move in the wage ladder. Jobs that are higher up in the ladder are typically more productive and, therefore, also more durable. Low skilled workers are known to be particularly affected by job loss. By contrast, high skilled workers typically hold relatively stable jobs. This pattern is also given here. The coefficient on the college dummy *Coll.* is highly significant and has a negative sign. A college graduate, ceteris paribus, has a 39% higher expected job duration than a worker without a college degree.

The last four coefficients reported in Table 3.2 measure how the duration of the preceding job affects the probability of separating from the current job. As mentioned previously, I include a quadratic term to account for possible non-linearities. In addition, I also condition on how the previous job ended. Let us first focus on the case $\bar{D} = 0$, i.e. the worker came into the new job after quitting the previous one. The linear and the quadratic term, expressed by τ_{prev} and τ_{prev}^2 are both highly significant. The estimates show that the duration of a worker's previous job positively affects the survival of the current one. One additional year of tenure (relative to the mean) raises the duration of the subsequent job by 16%. This effect is non-linear but decreasing, as indicated by the positive sign of τ_{prev}^2 .

The interaction term $\bar{D} * \tau_{prev}$ is also significant and has a negative sign. This is indeed interesting. It implies that a worker who is *fired* from a long term job has a higher expected duration in the new job than a worker who *quits* from the same job. This result, essentially, challenges the basic job ladder model. A central prediction of which is that long-term jobs ending in layoffs (quits) are, generally, followed by a job that is less (more) durable than the previous one. Here, instead, we find that new jobs after a layoff, ceteris paribus, are more stable than after a quit.

In the next step, I perform a series of hypothesis tests for the validity of the model specification. First, I test for the joint significance of the included explanatory variables. In particular, I estimate a model in which I consider only the baseline hazard rates, λ_n . This is achieved by setting $\eta = 0$. Then I compare the implied log likelihood with that of the unrestricted model. This is essentially a likelihood ratio test. The test statistic is asymptotically distributed as a χ^2 with 20 degrees of freedom. In our case, the test statistic is equal to 3,275.1, well above the critical value of 37.57 for the 99% confidence level with 20 degrees of freedom. Therefore, I can reject the restricted model. Next, I test the hypothesis that the coefficients of the age dummies are jointly equal to zero. The test statistic associated with this hypothesis is equal to 495.28, which leads to the rejection of the hypothesis at the 99% confidence level. At the same level of confidence, I also reject the hypothesis that all the coefficients on w are jointly equal to zero (test statistic: 186.22).

The question that remains to be answered is how well the predictions of the estimated model match up with the observed job durations in the data. To address this issue, I proceed in the following manner: For each observation j in the sample I compute the expected job duration predicted by the estimated model. This is done by using the observed characteristics \mathbf{x}_j and the expression in (3.9). The resulting predicted value is then compared with the actual duration of employment spell j . The discrepancy between both terms is denoted by $\phi = |t_{\text{predicted}} - t_{\text{actual}}|$.

I find that the mean absolute value of the prediction error, ϕ , is 0.97 years. Moreover, the proportion of jobs for which the prediction error is less than half a year is around 20%. For a maximum prediction error of one year, the proportion is 51% and for more than 90% of all spells the discrepancy ϕ is less than 3 years. The fit of the model increases substantially if we restrict attention to certain subgroups in the sample. Consider, for instance, the group of workers aged 25 years and below, with no college degree and who were laid off from their previous job. For 40.9% of all spells in this subsample, the prediction error ϕ is less than half a year. For 92.3% the discrepancy is less than one year. The large increase in the model's predictive power is partly driven by the lower average (actual) job duration in the sample, but also by the fact that certain subsamples, including the one considered here, are more homogenous in outcomes than others. This applies, in particular, to the group of young workers and those without a college degree.

3.4.2 Results for Question 2: Investigating the nature of job separations

This section reports the results obtained from estimating the Probit model in (3.5) - (3.7). Recall that the goal of the estimation is to identify factors that help to predict the nature of a separation. The dependent variable is $P(z_j = 1|\bar{\mathbf{x}}_j)$, i.e. the probability that job j ends due to a layoff, conditional on (a) a separation taking place and (b) a set of observables. The explanatory variables included in the estimation are the same as in the previous section. However, the specification differs slightly, as a worker's age (at match formation) and total tenure on the job are both treated as continuous variables. Moreover, I do not include the interaction term $\bar{D} * \tau_{prev}$, but instead, consider the worker's tenure on the previous job τ_{prev} and how he separated from it, \bar{D} , as two distinct variables⁸. Lastly, I also treat the number of weeks of non-employment \tilde{w} as a continuous variable. The interaction term with \bar{D} stays in place as before.

In total, 14 parameters are estimated: coefficients on (a) the constant term denoted c , (b) the worker's age age , age^2 , (c) the worker's total tenure τ , τ^2 , (d) the binary variable indicating the worker's education $Coll.$, (e) the difference between the initial and the last wage paid $\Delta\omega_{in}^{last}$, (f) the binary variable that indicates the nature of the previous separation \bar{D} , (g) the deviation of the worker's initial wage from the group average $\Delta\omega_{in.}$, (h) the wage gap between the current and the last job $\Delta\omega_{in}^{prev}$, (i) the worker's tenure on the previous job τ_{prev} , τ_{prev}^2 , (j) the duration of non-employment w , $D_{fir.*}w$. The coefficients are estimated by solving the maximization problem stated in (3.8). The results are reported in Table 3.3. In addition to the point estimates, I also report the change in the probability of facing a layoff, $P(z = 1|\cdot)$ induced by a change in each of the independent variables. The reference point for this comparison is the dismissal probability of a hypothetical average worker⁹.

The estimates reveal that the probability of experiencing a layoff is decreasing (at a decreasing rate) with age. Both coefficients, age and age^2 , are statistically significant. Notice again that "age" refers to a worker's age when he started the current job. The young are more likely to be fired than experienced workers. A rise in age by one year, relative to the average worker, lowers the probability of a layoff by 1.2%. This effect is decreasing with age as indicated by the positive sign on age^2 . It reached its maximum at the age of 45 years and

⁸Preliminary results reveal that the interaction term $\bar{D} * \tau_{prev}$ is insignificant, also when a quadratic term is added.

⁹The average worker in the sample faces a layoff probability of 27.87%, conditional on experiencing a separation.

declines thereafter.

Coeff.	Value	$\Delta_{av.}$	$\Delta_{ref.}$	Coeff.	Value	$\Delta_{av.}$	$\Delta_{ref.}$
<i>const.</i>	0.208 (0.203)			\bar{D}	0.426 (0.032)	40.7	62.1
<i>age</i>	-0.049 (0.015)	-11.7 (1 yr.)		$\Delta\omega_{in}$	0.232 (0.043)	0.8 (+10%)	
<i>age</i> ²	0.0007 (0.0002)			$\Delta\omega_{in}^{prev}$	0.063 (0.034)	2.8 (+10%)	
τ	-0.043 (0.011)	-3.9 (1 yr.)		τ_{prev}	-0.025 (0.012)	-2.2 (1 yr.)	
τ^2	0.003 (0.001)			τ_{prev}^2	0.002 (0.001)		
<i>Coll.</i>	-0.337 (0.028)	-28.7	-34.9	\tilde{w}	0.003 (0.001)	1.9 (1 wk.)	
$\Delta\omega_{in}^{last}$	0.128 (0.058)	1.5 (+10%)		$\bar{D} * \tilde{w}$	-0.004 (0.002)	-0.2 (1 wk.)	-4.5 (1 wk.)

Dependent variable: $y = Prob(layoff|x_i, separation)$, Mean: $E(y) = 0.2787$
Obs.: 14,835, $\log L = -8,792$, Percentage correctly predicted: 70.02%
LR Test: H_0 : only *constant* (p: 0.000)

The coefficients are labeled in the same way as the respective covariate. $\Delta_{av.}$ and $\Delta_{ref.}$ are defined as in Table 3.2. Standard errors are in parentheses.

Table 3.3 Estimation Results: Probit Model

The same pattern applies to a worker's total tenure on the current job. Short term jobs are more likely to end due to a layoff than long term jobs. One more year of tenure lowers the probability of a layoff by 3.9%. This effect reaches its maximum at a duration of 8.5 years and gradually diminishes for $\tau > 8.5$. Education guarantees more stable jobs, as shown in the previous section, and also shields workers from layoffs. The coefficient on the college dummy is negative and highly significant. The estimate implies that college graduates are 34.9% less likely to be fired than workers without a degree.

Next, let us next focus on the effect of wages on the layoff probability. The coefficient on $\Delta\omega_{in}^{last}$ is positive, indicating that jobs with positive wage growth are more likely to end with a layoff than jobs with a flat wage profile. A 10% wage gain on the job (relative to the mean) raises the layoff probability by 1.5% implying that, economically, this effect is not too important. The positive sign is intuitive: Positive wage growth raises the value of the worker's current job and makes it is less likely that the worker finds it optimal to quit and leave the firm for a better job. Since $P(z = 1|\cdot) = 1 - P(z = 0|\cdot)$ and $\partial P(z = 0|\cdot)/\partial\omega_{in}^{last} < 0$ we observe a rise in the layoff probability.

A similar logic can be applied to interpret the coefficient on $\Delta\omega_{in}^{prev}$, which is positive. The larger the wage gain of moving to a new job, the lower probability that the worker quits. Hence, when we observe a separation it is more likely to be due to a layoff. The estimates imply that a 10% wage gain between jobs translates into an increase in the layoff probability by 2.8%. A worker who jumps upwards in the wage ladder is less likely to quit than a worker that falls downwards. Consequently, we expect the former to have a higher relative probability of being fired than the latter. Further, the coefficient on $\Delta\omega_{in}$ is positive and in contrast to the previous two, highly statistically significant. However, the deviation of the initial wage from the group average appears to be of only minor economic importance only. A positive deviation of 10% raises the relative layoff probability by just 0.8%, which is rather negligible. A more important consideration is how a worker's preceding job has ended.

The coefficient on the dummy variable \bar{D} is positive and highly significant. The value of the estimate implies that a previously laid-off worker is 62.1% more likely to be laid off from the current job than a worker who previously quit. Two possible reasons for this pattern could be: (1) The existence of a path dependency in the labor market as a dismissal may stigmatise a worker for his future career path, or (2) There may be an inherent difference in turnover across workers. Some workers just change jobs more often than others do. Duernecker (2009) shows that the career-total number of jobs differs enormously across workers. This may be an indication for such inherent differences that cannot be explained by the observable job and worker characteristics used here. In the current framework, one could control for the job turnover rate at the individual level and this may shed some light on the issue.

One more year of tenure on the previous job lowers a worker's current layoff probability by 2.2%. This effect might be due to the accumulation of skills that are (at least partly) transferable across jobs. The longer a job lasts, the more skills a worker can accumulate and, therefore, the more valuable he will be for any subsequent employer. Thus, the risk of being fired is lower.

The waiting time until reemployment increases the layoff probability in the new job. Workers who stay out of employment longer are more likely to be fired than those that return more quickly. An additional week raises the relative layoff probability by 1.9%. An explanation for this pattern could be the depreciation of skills during a spell of unemployment. The longer the spell, the more skills are lost and the less valuable a worker is for the subsequent employer. The coefficient on the interaction term $\bar{D} * \tilde{w}$ is negative but its significance is rather weak. Hence, on a formal basis, we cannot reject the null that the effect of \tilde{w} on the relative layoff probability is the same after a quit and after a layoff.

Lastly, I perform a likelihood ratio test to test for the joint significance of the covariates. To this end, I first estimate a restricted model that has only a constant term and then compare the implied log likelihood with that of the unrestricted model. The test statistic, asymptotically distributed as a χ^2 with 13 degrees of freedom, is equal to 560.56 and is well above the critical value of 27.69 for the 99% confidence level with 13 degrees of freedom. Therefore, the restricted model is rejected. Moreover, I test for the joint significance of all the variables that appear (a) in levels and squared or (b) together with an interaction term. The hypotheses that are tested are the following (1) $age = 0, age^2 = 0$, (2) $\tau = 0, \tau^2 = 0$, (3) $\tau_{prev} = 0, \tau_{prev}^2 = 0$ and (4) $\tilde{w} = 0, \bar{D} * \tilde{w} = 0$. The test statistics for hypotheses (1) and (2) are, respectively, 17.75 and 14.31, which are above the critical value of 9.21 for the 99% confidence level with 2 degrees of freedom. Thus, we can reject the null that age and tenure have no significant effect on the layoff probability. The test statistic for (3) is 3.97 and implies that we can not reject the third hypothesis, even at the 90% confidence level. Hypothesis (4) is rejected only at the 99% level but not at the 95% level of confidence. Lastly, I evaluate the predictive power of the model. To that end, I compute the fraction of separations in the sample for which the model correctly predicts the nature. I find that 70.02% of all observations are correctly predicted.

3.4.3 Results for Question 3: Reemployment Wages

This section closes the circle. The first questions addressed ongoing employment spells and investigated the factors influencing the likelihood of job separation. The second question

took over at the point when a separation occurred and aimed at identifying factors that help predict whether employment spells will end due to a layoff or a quit. After the separation from a job, individuals eventually become reemployed in a new job. This is the point where the third question becomes relevant. It considers the wage of the previous and of the new job (the last and the initial wage respectively) and investigates how much a worker gains or loses by transiting to a new job. For most of the analysis, I will condition on how the previous job ended. More specifically, I compute wage gaps separately for jobs that follow a layoff and those that follow a quit. Quits and layoffs inherently differ from each other not only in cause but also in the implications, as shown in the previous two sections. Thus, it is natural to consider both cases separately.

The variable of interest in this section is $\Delta\omega_{in}^{prev}$, i.e. the wage differential between the current and the last job. More precisely, $\Delta\omega_{in}^{prev}$ is the difference (in percentage terms) between the initial wage paid to a worker in job j and the last wage paid in the job preceding spell j . What are the determining factors and magnitude of the "between-jobs" wage differential? These are the questions that will be the main focus of this section. Unlike the previous two sections, the issue of this section is not addressed within a formal econometric framework. The reason for this is the enormous dispersion of $\Delta\omega_{in}^{prev}$ in the sample, even within narrowly defined groups. The significance of any point estimate obtained from an econometric model would, most likely, be eroded by high standard errors¹⁰. For the moment, however, I refrain from performing a fully fledged econometric analysis. Instead, I report a set of descriptive statistics that are meant to provide just a rough illustration of the issue. I will focus on non-college worker only, mainly for homogeneity reasons. For comparison, I often report also the respective figure for college workers. Moreover, the main statistic of interest will be the sample-median of $\Delta\omega_{in}^{prev}$. The high dispersion makes the median the more informative measure than, for instance, the mean as it also mitigates the bias induced by extreme values.

The median wage gain associated with moving to a new job, coming either directly from another job or via a spell of unemployment, for U.S. white male non-college workers aged 18 years and above is 3.72% (college workers: 6.73%). After a quit the median wage **gain** is 5.6% (8.48%) and after a layoff the median wage **loss** is 0.05% (0%). For a more differentiated picture, I split the sample into different age groups and compute $\Delta\omega_{in}^{prev}$ for each group separately. The results are reported in Table 3.4. The first column depicts the outcomes for

¹⁰Though controlling for unobserved heterogeneity in the sample might help to get reasonable estimates.

four different age groups, namely $[18 - 23)$, $[23 - 28)$, $[28 - 33)$ and $[33 - 40]$. The wage gains after a quit are relatively homogenous across age groups. They are the highest for workers aged 18 – 23 years, 6.92% (11.08%), and are slightly lower for higher age groups. The gain is lowest for age group 28 – 33 years, 4.93% (9.5%). Interestingly, for workers aged between 33 – 40, the wage gain starts rising again to a rate of 5.17% (5.73%). Two points are worth mentioning here: (1) Over the life cycle until the age of 40, the wage gains for non-college workers are U-shaped. As we will see later on, this pattern is remarkably stable even when we consider more refined groups. (2) The wage gains are generally higher for college-educated workers. For this group, the gains do not feature a U-shape but are strictly decreasing with age.

	All		$\tau_{prev} \geq 1$ yr.		$\tau_{prev} \geq 2$ yrs.		$\tilde{w} \geq 5$ wks.	
	Quit	Firing	Quit	Firing	Quit	Firing	Quit	Firing
All	5.79	-0.05	6.06	-2.96	4.21	-3.32	3.52	-0.14
18-23	6.92	1.94	12.71	0.61	15.18	2.62	4.93	4.47
23-28	5.06	-0.02	6.32	-5.25	4.99	-7.74	3.52	-1.12
28-33	4.93	-1.88	2.50	-9.33	0.00	-10.20	2.06	-5.74
33-40	5.17	-0.09	5.97	-0.05	3.85	0.00	0.77	-1.33

Median wage gap in percent between previous and current employment spell. τ_{prev} is the duration of the previous job, \tilde{w} is the duration (in weeks) until reemployment. Sample: U.S. white male workers. Data: *NLSY* 79.

Table 3.4 Reemployment Wages

Let us next turn to post-layoff reemployment wages. The median wage gap is generally lower than after a quit, as would be expected, and the differences across age groups are more pronounced. The second column in Table 3.4 shows that young workers generally gain from being fired (in terms of wages) but prime-age workers loose. The median wage gap for the group 18 – 23 is positive and equal to 1.94% (4.84%) and negative for workers older than 23 years. The largest firing losses, 1.88%, are experienced by the group aged 28 – 33. The U-shape is readily identifiable also here.

Why do the young gain from a firing but prime-age workers do not? Specific human capital might be an explanation as the wage earned in an occupation is likely to be based partly on the worker's level of specific skills. Prime-age workers typically have higher job tenure than the young. Thus, they also accumulate more (firm) specific skills. In case of a layoff, those skills are destroyed and reemployment wage are lower. By contrast, the young have less specific skills to loose. In addition, at the age of 18 – 23 years, individuals are still in the process of finding the right match. Hence, a layoff will often end a bad match after which the worker continues the search process. Any new job offer (that is accepted) is likely to be a better paid job.

The duration of the previous job is obviously an important factor. To further investigate this issue I now consider all those jobs for which $\tau_{prev} \geq 1$ year and $\tau_{prev} \geq 2$ years, i.e I consider jobs for which the preceding employment spell lasted longer than one year and two years. When considering the full sample, I find that the gains for a quit remain positive and substantial. For jobs with $\tau_{prev} \geq 1$ the median gain is 6.1% and it is slightly lower for $\tau_{prev} \geq 2$, namely 4.2%. For layoffs there is a clear pattern emerging as we increase τ_{prev} . In the pooled sample the firing losses are negative (−0.05%) but negligible. The longer is the duration of the previous job the more individuals lose in terms of reemployment wages. For jobs with $\tau_{prev} \geq 1$ and $\tau_{prev} \geq 2$ the median loss, respectively, is 2.9% and 3.3%. This is consistent with the existence of specific skills. Longer-lasting jobs are associated with more specific-skills that are destroyed in the event of a layoff. Notice that for both quits and layoffs, the U-shape is preserved across age groups. Young workers are harmed least by a layoff. In all cases depicted here, the firing gains for the age group 18 – 23 years are positive and clearly higher than for any other groups. By contrast, prime age workers generally suffer a substantial loss from a layoff. For those aged 28 – 33 it amounts to a loss of 9.3% and even 10.2% when $\tau_{prev} \geq 2$.

The last case depicted in Table 3.4 reports the median wage differential for jobs that began after a spell of non-employment of at least 5 weeks. In general, the quit gains are lower and the firing losses are more negative than in the full sample. Evidently, the duration until reemployment worsens individuals' wage prospects. One thing is especially worth looking at: Arguably, a quit that is followed by more than 5 weeks of non-employment hardly reflects a job-to-job transition. A quit of this sort is probably motivated by things other than an imminent job switch. Therefore, an obvious question is, what distinguishes two individuals each with $\tilde{w} \geq 5$ where one has quit the previous job while the other was laid-off? This

question is especially relevant in light of the fact that the average duration of \tilde{w} , conditional on $\tilde{w} \geq 5$, is almost the same after a quit (20.37 weeks) as after a layoff (20.58 weeks). Despite a \tilde{w} that is roughly the same, we observe a wage gap that is very different for quits and firings. This pattern generally holds for the prime age groups but strikingly, however, for the young the difference is practically nonexistent.

3.5 Discussion and Conclusion

The purpose of this paper is threefold: (1) to identify important determinants of job separation, (2) to find relevant factors that help predict whether a given employer-employee match will be dissolved due to a quit or a layoff and (3) to quantify the respective effects of a quit and a layoff on reemployment wages. These issues are addressed using data on white males from the National Longitudinal Survey of Youth (1979). The main results of the paper can be summarized as follows:

Job tenure significantly reduces the likelihood of separating from the current employer. This is a well known result and many studies in the literature find that the separation hazard exhibits a negative duration dependence. New in this paper is that job tenure reduces the hazard not only in the current but also in the worker's subsequent employment spell. The longer a worker is employed in one job, the more durable his next one will be. The estimates imply that a one-year increase in current tenure raises the subsequent job duration by 16 – 25%. Very importantly, this result is independent of whether the current match separates due to a quit or a layoff. The negative duration dependence after a layoff is indeed a surprising result and challenges the standard job ladder model. A key prediction of which is that after a layoff, short **and** long duration jobs are both followed by a series of short-term jobs. This prediction is contradicted by the data as it violates the negative duration dependence across subsequent jobs.

Current tenure improves future (and current) job stability but it generally lowers reemployment wages after a job change. This is especially true in the case of a layoff. Individuals that are being laid off from a medium term job suffer significantly higher wage losses than workers that are displaced from short term jobs. An explanation for this is likely to be the destruction of (firm) specific skills in the event of a layoff.

A quit and a layoff have markedly different consequences for a worker's subsequent employment outcomes. After a layoff, a worker typically holds relatively short and unstable jobs that are likely to result in another layoff. Moreover, the wage loss upon reemployment

can be substantial. By contrast, a quit often reflects a transition to new, and also stable, job which is accompanied by a wage gain. With respect to quit gains and firing losses we observe significant differences across age groups. After a layoff, the young enjoy higher reemployment wages whereas all other workers generally suffer a wage loss. The gains after a quit are positive for all age groups but they are again highest for the young.

The typical pattern is that the young hold rather unstable jobs and are more likely to be laid off than prime age workers. However, each time they separate and transit to a new job they, generally, experience a wage increase. By contrast, prime age workers seem to hold more stable jobs and any separation is more likely to be due to a quit than a layoff. However, if they are laid off they generally suffer a substantial wage loss. From a general point of view, job loss and instability for the young are not a bad thing after all as they both seem to promote the efficient reallocation of labor. In contrast, job loss for prime-aged workers can have serious negative, possibly long-lasting, effects on their career paths.

The spell of non-employment between two consecutive jobs is an important determinant of a worker's prospects after reemployment. Generally speaking, the longer the period between job separation and reemployment, the more workers lose in terms of wages. An explanation could be the loss of skills during non-employment. At the same time, we observe that jobs that started after a long non-employment spell are more likely to result in a layoff than in a quit. Furthermore, I find that the returns to job search, in terms of future job stability, are positive and substantial. A job that is accepted very quickly, within 0-2 weeks after the layoff, is likely to last considerably shorter than a job that is selected within 3-10 weeks. This result lends support to the hypothesis presented in Marimon and Zilibotti (1999), that workers are more likely to end up in a mismatch when they take up a new job relatively quickly.

Several limitations should be considered when interpreting the results of this paper. The specifications estimated in Section 3.2 do not control for business cycles or industry fixed effects, both of which may be important. Job separation rates are typically pro-cyclical as matches are more likely to dissolve during recessions. Hence, it would make sense to include a variable in the estimation that indicates the current, aggregate state of the economy. Omitting such a variable could cause a bias in the estimates. Industry fixed effects are due to inherent differences across industries in worker turnover rates. Some industries, such as many of those within the service sector, exhibit substantially larger job separation rates than others. By not considering these differences we run the risk of mistakenly attributing a substantial part of the observed separations to the explanatory variables. Future research will address these concerns.

Bibliography

- [1] Burdett, K. and M. Coles (2003). Equilibrium wage-tenure contracts. *Econometrica* 71 (5), 1377-1404.
- [2] Burdett, K. and D. Mortensen (1998). Equilibrium wage differentials and employer size. *International Economic Review* 39 (2), 257-274.
- [3] Duernecker, G. (2009). Informational Frictions and the Life-Cycle Dynamics of Job Mobility, University of Mannheim mimeo.
- [4] Fallick, B. and C. Fleischman (2004). The importance of employer-to-employer flows in the U.S. labor market. Federal Reserve Bank Board of Governors mimeo.
- [5] Marimon, R. and F. Zilibotti (1999). Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits. *Economic Journal* 109 (455), 266-91
- [6] Pissarides, C. A. (1994). Search unemployment with on-the-job search. *Review of Economic Studies* 61 (3), 457-75.