

Learning, Mobility, and Wage Dynamics: Theory and Evidence

Uta Schönberg

A dissertation submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy
University College London**

February 2004

ProQuest Number: U643902

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest U643902

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

There are (at least) three benchmark models, or 'building blocks', for explaining labour market outcomes: human capital models, search and matching models, and learning models. The objective of this thesis is to assess the empirical importance of those models.

Chapter 2 focuses on learning models. The question addressed in this chapter is: Is employer learning symmetric or asymmetric? A model that nests both learning hypotheses is developed. Tests that allow to discriminate between the two hypotheses are proposed. Evidence from the National Longitudinal Survey of Youth indicates that asymmetric learning plays an important role for university graduates, but not for high school graduates and dropouts.

Chapter 3 focuses on human capital models. It revisits the question why firms pay for general training within the German apprenticeship system. The focus is on the impact of wage rigidities caused by unions on training. A model of firm-sponsored training and unions is developed, and empirical implications are derived. The empirical evidence suggests that wage floors created by unions are an important reason for firm-financed training in Germany. However, asymmetric employer learning cannot be ruled out as an additional reason for firm-financed training.

Chapter 4 focuses on search and matching models. It compares job mobility of young men in two countries with very different labour market institutions, the United States and Germany. Match-specific productivity plays an important role in both countries for all education groups. The proportion of the variance of the unexplained component of wages that can be attributed to match quality is considerably higher in Germany than in the US. In Germany, there is evidence that job search and match-specific productivity matters more for unskilled workers than for apprentices.

Contents

List of Tables	6
List of Figures	8
Acknowledgements	9
1 Introduction	10
2 Testing for Asymmetric Employer Learning	19
2.1 Existing tests for employer learning	23
2.2 The learning model	25
2.2.1 Description of environment	25
2.2.2 Symmetric learning	29
2.2.3 Asymmetric learning	31
2.3 Empirical implications	32
2.3.1 Stayers, job-to-job and job-to-unemployment movers	33
2.3.2 Modification of existing tests for symmetric learning: Learning by experience or tenure?	35
2.3.3 Test for the single crossing assumption	37
2.3.4 Summary and discussion of assumptions	37
2.4 Data description and variables	41
2.5 Empirical evidence	43
2.5.1 Stayers, job-to-job, and job-to-unemployment movers	43
2.5.2 Learning by experience or tenure?	47

2.5.3	Alternative explanations	51
2.6	Conclusion	55
2.7	Appendix	58
3	Training and Unions¹	67
3.1	Collective bargaining in Germany and outline of model	71
3.2	A model of firm-financed training	72
3.2.1	Base model	73
3.2.2	Union agreements	77
3.3	Empirical Implications	85
3.3.1	Training in unionised and non-unionised firms	85
3.3.2	Worker selection: Wages of movers and stayers in unionised and non-unionised firms	88
3.4	Empirical Analysis	92
3.4.1	Data Sources	92
3.4.2	Training in unionised and non-unionised firms	92
3.4.3	Worker selection	95
3.5	Discussion and Conclusion	101
3.6	Appendix	104
4	Job Mobility in the United States and Germany	120
4.1	A model of job-to-job and job-to- unemployment mobility	123
4.1.1	Description of environment	123
4.1.2	The worker's mobility decision	127
4.2	Empirical implications	131
4.2.1	Wage gains by mobility groups	131
4.2.2	The covariance structure of log-wage residuals	133
4.3	Data description and variables	141
4.3.1	German data	141
4.3.2	US data	142

¹This chapter is joined work with Christian Dustmann.

4.4	Empirical evidence	144
4.4.1	Job-to-job and job-to-unemployment mobility	144
4.4.2	Wage gains by mobility groups	147
4.4.3	The covariance structure of log-wage residuals	148
4.5	Summary	160
4.6	Appendix	163
5	Concluding Remarks	167
	Bibliography	170

List of Tables

1.1	Labour market dynamics: United States versus Germany	11
2.1	Specification of workers' productivity	26
2.2	Empirical implications of symmetric and asymmetric learning	38
2.3	Means and standard deviations of selected variables	43
2.4	Average AFQT score of job-to-job and job-to-unemployment movers . .	45
2.5	Learning by experience or tenure?	47
2.6	Learning by experience or tenure? Results by education	50
2.7	Average AFQT score of job-to-job and job-to-unemployment movers: The effect of training measures	52
2.8	Learning by experience or tenure? The effect of training measures . . .	54
3.1	Unionised and non-unionised firms	93
3.2	Effect of Unionisation on Training Intensity	95
3.3	Wages of stayers, endogenous and exogenous movers	97
3.4	Mover, fixed training firm effect estimates	98
3.5	Wages of movers and stayers: Unionised versus non-unionised firms . .	100
3.6	Military and Social Service: 3 Definitions	116
3.7	Variable definitions: Firm data	118
4.1	The autocovariance of log-wage residuals by mobility groups	149
4.2	The variance of ability by experience: United States	151
4.3	The variance of ability by experience: Germany	152
4.4	Autocorrelation of the within-job wage growth residual	153

4.5	The variance of the within-job wage growth residual by tenure: United States	154
4.6	The variance of the within-job wage growth residual by tenure: Germany	155
4.7	The variance of the log-wage residual and match quality by experience: United States	158
4.8	The variance of the log-wage residual and match quality by experience: Germany	159
4.9	The variance of ability and match quality by experience: United States	160
4.10	The variance of ability and match quality by experience: Germany . . .	160
4.11	Means and standard deviations of selected variables: US and German data	163
4.12	Wage censoring by experience and education: German data	164

List of Figures

3.1	Wage determination in unionised firms	79
3.2	Socially optimal training and training in unionised firms	83
3.3	Wages of stayers, endogenous and exogenous movers by experience . . .	99
3.4	Profit with and without training as a function of workers expected ability	106
3.5	Difference between the utility from working in a unionised and non- unionised firm	109
4.1	The variance of within-job wage growth residuals by tenure	140
4.2	Number of jobs by experience and education	145
4.3	Quarterly empirical job survival rate by education	145
4.4	The proportion of job-to-job transitions by experience: United States versus Germany	146
4.5	Within-job, job-to-job and job-to-unemployment wage growth: United States versus Germany	148

Acknowledgments

I would first and foremost like to express my gratitude to my supervisors Christian Dustmann and Costas Meghir. They have taught me a lot. This work would have been impossible without them. I have also greatly benefited from discussions with my external examiners, Joe Altonji and Chris Pissarides.

I have received many fruitful comments and suggestions from Richard Blundell, Kevin Lang, Derek Neal. I also would also like to thank Arndt Kölling, Holger Alda and in particular Stefan Bender from the IAB for their invaluable help with the German data.

Finally, I would like to thank some of my colleagues and friends at UCL who shared many of my research interests. I would like to cite, amongst many others, Chiara Criscuolo, Fernando Galindo-Rueda and Peter Postl. The usual disclaimer applies: all remaining errors are my own.

This work would not have been possible without the financial support received from the Economic and Social Research Council and the Gottlieb Daimler and Karl Benz Foundation. I am very grateful to them.

Chapter 1

Introduction

Early labour market careers are characterised by many job switches. In a sample of young men from the National Longitudinal Survey of Youth (NLSY79), the typical worker has switched jobs about 5 times in ten years (table 1.1, column 1)¹. For comparison, the typical German worker is on his third job after ten years (table 1.1, column 2). Over the same period his wage grows substantially. In both countries the average wage growth is about 72 %². Despite the high average growth rate, there is a considerable fraction of workers in the US whose wage after ten years in the labour market is no higher than at labour market entry³. In Germany, in contrast, this fraction is negligible.

There are three benchmark models, or 'building blocks', for explaining labour market outcomes of workers: human capital models, search and matching models, and learning models. The human capital theory dates back to Becker (1964). It is still the main theory for wage growth over the life cycle. Becker (1964) distinguishes between

¹Topel and Ward (1992) find that the typical worker in the US holds seven jobs during the first ten years in the labour market, about 2/3 of his life time total.

²According to Topel and Ward (1992) the worker's wage almost doubles.

³Baker et al. (1994a, 1994b) report similar results using very different data. Their data contains confidential personnel records for all management employees of one large firm (more than 60.000 employees) in a service industry over the years 1969-1988. Management workers constitute 20 % of the total work force of the firm. They find that in a typical cohort of entrants, about 15 % experience a real wage decline over a period of ten years.

Table 1.1: Labour market dynamics: United States versus Germany

	United States	Germany
number of jobs after ten years	5.66 (3.45)	2.82 (1.95)
cumulative real wage growth in ten years ($\ln w_{10} - \ln w_0$)*	0.54 (0.506)	0.54 (0.362)
porportion of workers whose real wage after ten years is no higher than at labour market entry	0.13	0.02
proportion of workers whose real wage does not increase in any year	0.07	0.01

* This implies a wage growth of $e^{0.54} - 1 = 0.72$. US data: NLSY79. See chapter 2, section 2.4 for a description of the data and sample selection. N=1720. German data: IAB-Beschäftigtenstichprobe. See chapter 4, section 4.3 for a description of the data and sample selection. N=5180. Standard deviation in parenthesis.

general and firm-specific human capital⁴. General human capital refers to training that is equally valued in many firms, while firm-specific human capital refers to training that is only valuable at the firm that provides training. While the return to firm-specific human capital is still in dispute⁵, there is consistent evidence for considerable returns to general human capital⁶.

Burdett (1978), and Jovanovic (1979a, 1979b) are examples of search and matching models. Search and matching models assume that the worker's productivity is match-specific, i.e. his productivity in one firm is uncorrelated with his productivity in another firm. Workers do not know the location of the firm at which they are best matched, and search for a good match. Some of the empirical implications of search models depend on whether the job is an 'inspection' or 'experience' good⁷. A job is a pure inspection good if the worker's productivity becomes fully known as soon as the firm and worker meet. If, on the other hand, new information about the job arrives after the worker has accepted the job and 'experiences' it, the job is -partly- an experience

⁴Recently, Neal (1995) and Parent (2000) have stressed the importance of industry-specific human capital. Industry-specific human capital is portable to firms within the same industry, but not to firms in other industries.

⁵For the US, Topel (1991) finds a cumulative return to tenure of 0.21 in ten years. Altonji and Shakotko (1987) and Abraham and Farber (1987), in contrast, find only modest returns to tenure.

⁶See also Altonji and Williams (1998).

⁷This terminology goes back to Nelson (1970).

good. Since on-the-job search allows workers to allocate better matches with time in the labour market, search models provide an additional explanation for why wages grow with time in the labour market. For the US, Topel and Ward (1992) find that job mobility is an important contributor to wage growth: Their estimates imply that one third of overall wage growth in the first ten years can be attributed to job search.

Early examples for learning models are Waldman (1984), Greenwald (1986) and Harris and Holmstrom (1982)⁸. The intuition behind learning models is simple. When workers enter the labour market, firms are only imperfectly informed about their productivity. The longer a job-candidate's employment history, the more information a recruiting firm can draw upon in assessing the candidate's ability to do the job he is being considered for. Learning models can explain why some workers do not experience a real wage increase over a long time period although on average wages substantially rise. Search models in which the job is an experience good also contain a learning element. There is, however, a crucial difference. In contrast to search models, learning models -or models which are termed 'learning models' in this dissertation- assume that the worker's productivity is general, i.e. his productivity in one firm is perfectly correlated with his productivity in another firm. If learning is about general as opposed to match-specific productivity, the distinction between symmetric and asymmetric learning becomes important. Symmetric learning means that incumbent and outside firms have the same information about the worker's productivity. Asymmetric learning, in contrast, means that recruiting firms have an informational disadvantage relative to workers' current employers. Both learning hypotheses have found empirical support⁹. The distinction between symmetric and asymmetric learning is irrelevant if learning is about match-specific productivity, as the information of incumbent firms is of no value

⁸In Waldman (1984) and Greenwald (1986) the learning process is asymmetric. Waldman (1984) is concerned with task assignment within the firm. Greenwald (1986) analyses the consequences of asymmetric learning for turnover. In Harris and Holmstrom (1982) the learning process is symmetric. They consider risk-averse workers and focus on insurance aspects.

⁹Farber and Gibbons (1996) and Altonji and Pierret (2001) derive empirical implications and test for symmetric employer learning. Gibbons and Katz (1991) derive empirical implications and test for asymmetric employer learning.

to outside firms.

The objective of this thesis is to assess the empirical importance of each building block model. I look at two countries with very different labour market institutions, the United States and (West) Germany. The data for the US comes from the 1979 National Longitudinal Survey of Youth (NLSY79). The empirical analysis for Germany is based on a new administrative data set, the so-called *IAB-Beschäftigtenstichprobe*. Both data sets allow to construct an accurate and complete work history of workers from labour market entry onwards, and are thus well suited for the analysis of job mobility and wage dynamics. Chapter 3 additionally uses a new and unique data set that combines survey data on firms, the so-called *IAB-Betriebspanel*, with administrative data on workers. The data was provided by the Institute for Labour Market Research (IAB) in Nuremberg, Germany.

Chapter 2 focuses on learning models. The question addressed in this chapter is: Is employer learning symmetric or asymmetric? A model that nests both learning hypotheses is developed. It is shown that existing tests that have been derived under one learning hypothesis are also consistent with the other. New tests that allow to discriminate between the two hypotheses are proposed. A key difference between my and most other learning models is endogenous mobility: In my model, there are endogenous job-to-job as well as job-to-unemployment transitions. Mobility arises because workers do not only care about wages, but also about non-pecuniary characteristics, such as distance from work, health plans, relationship with co-workers, etc.

The testing strategy proceeds in two steps. Both tests make use of a variable that is an indicator for workers' ability but not observed by employers, such as the AFQT score. The first test compares three groups of workers: stayers, job-to-job and job-to-unemployment movers. I show that under symmetric learning ability has no impact on the probability of moving from job-to-job. Under asymmetric learning, in contrast, ability negatively affects the probability of a job-to-job movement. Consequently, job-to-job movers are as able as stayers under symmetric learning, and less able than stayers under asymmetric learning. Both learning hypotheses are consistent with a lower ability of job-to-unemployment movers. I test these implications by comparing

the AFQT score of the three groups of workers.

The second test is a modification of Farber and Gibbons' (1996) and Altonji and Pierret's (2001) test for symmetric learning. Symmetric and asymmetric learning differ with respect to how a hard- and easy-to-observe variable, such as ability and education, affect wages for those who stay and move. Under symmetric learning incumbent and outside firms have the same information about workers' ability and education. Hence, the AFQT score and education affect wages for movers in the same way as wages for stayers. In contrast, under asymmetric learning incumbent firms use the newly acquired information about workers' ability, while outside firms continue to use education as a signal for workers' ability when determining wages. The coefficient on the test score is thus higher for stayers than for movers, while the coefficient on education is higher for movers than for stayers. Using data from the NLSY79, I find evidence for asymmetric learning for university graduates, but not for high school graduates and dropouts.

Chapter 3 focuses on human capital models. This chapter is joint work with Christian Dustmann. Standard human capital theory predicts that in perfectly competitive labour markets firms do not invest in general training of their workers, as the worker captures the full return to that investment (Becker (1964)). Yet, there is evidence that suggests that firms provide and pay for general training. A prominent example is the German apprenticeship system. The apprenticeship system is a vocational education scheme that combines firm and state provided training¹⁰. A crucial feature of the system is that - despite training workers in mostly general skills - it is, at least partly, financed by firms¹¹.

In recent work, Acemoglu and Pischke (1999a, 1999b) offer an explanation for this puzzle. The key to their argument is a compressed wage structure: If training increases workers' productivity by more than workers' outside option, then firms can increase profits by training. Firms thus have an incentive to bear the cost of training - even if training is completely general. One reason for a compressed wage structure

¹⁰Details on the apprenticeship can for instance be found in Soskice (1994), Franz and Soskice (1995) and Winkelmann (1997).

¹¹See studies by von Bardeleben et al. (1995) and Acemoglu and Pischke (1999b) for evidence.

is asymmetric employer learning (Acemoglu and Pischke (1998)). Other reasons are labour market institutions, such as unions and a minimum wage.

The third chapter revisits the question why firms pay for general training within the German apprenticeship system. The focus is on the impact of wage rigidities caused by unions on training. In Germany unions play an important role: Between 1996 and 1999, 77 % of the West German work force were covered by collective bargaining agreements¹². Our empirical test exploits that in Germany only firms that belong to an employer association are bound to union agreements. In firms that are members of an employer association union agreements apply to all workers, regardless of the worker's union status. This essentially creates a unionised and non-unionised sector¹³.

We first develop a model of firm-sponsored training and unions. The key feature of the model is that a unionised and non-unionised sector co-exist. We derive wage and training determination in both sectors, and explicitly analyse the worker's decision in which sector to work. The firm's training decision is analysed under the assumption that firms cannot fully commit to training provision. Our main theoretical results can be summarised as follows. First, an adverse selection of movers -the key implication of Acemoglu and Pischke's asymmetric learning model- may also follow from our union model. Second, unionised firms offer more training than non-unionised firms. If firms cannot fully commit to training provision, unions move training closer to the socially optimal level. Third, the impact of unions on training is not uniform. Workers sort into the unionised sector based on the impact union agreements have on training. Furthermore, more able workers also self-select into the unionised sector. Finally, long-term wage contracts mitigate, though not eliminate, the problem of market failure due to limited commitment to training provision. Such contracts, however, are not self-enforceable. The role unions essentially play in our model is that they make long-term wage contracts enforceable.

Our empirical strategy proceeds in two steps. In a first step we test for an adverse selection of movers. In a second and key step we use matched employer-employee

¹²The number is based on the IAB-Betriebspanel. See chapter 3, table 3.1.

¹³See chapter 3, appendix A for a description of the German collective bargaining system.

data to directly test whether unions increase training. Our empirical analysis attempts to take into account a nonrandom selection of workers and possibly firms into the unionised sector. We conclude that wage floors created by unions are an important reason for apprenticeship training in Germany. However, we cannot rule out asymmetric information as an additional reason for apprenticeship training.

Chapter 4 focuses on search and matching models. It compares job mobility of young men in two countries with very different labour market institutions, the United States and West Germany. It is well known that German workers switch jobs less often than American workers¹⁴. There exist opposing views about the difference in mobility rates. On the one hand, the lower mobility in Germany is regarded as positive. Proponents of this view stress the wastefulness of the search process. In this context the German apprenticeship system is sometimes praised. It has been credited with smoothing the transition from school to work and lowering youth unemployment (e.g. Ryan (2001)). Others, on the other hand, emphasise the importance of job search for wage growth and allocative efficiency, and regard the lower mobility in Germany as negative. Proponents of this view sometimes see the apprenticeship as an obstacle to job mobility. For instance, referring to the German apprenticeship system, Neal (1999) writes "institutions that limit returns to search may lead to ... an inefficient assignment of workers to tasks in the economy" (p. 257).

The fourth chapter provides new evidence on the job search process in both countries. I begin with the set-up of a search model that guides the empirical analysis. The search model borrows from Jovanovic (1984) and Mortensen (1988). The two crucial features of the model are on-the-job search and learning about match quality. On-the-job search leads to endogenous job-to-job transitions, and allows workers to allocate better matches over time. Learning about match quality leads to endogenous job-to-unemployment transitions: A match that used to be preferable to unemployment may not be preferable anymore after disappointing news about the match arrived. A

¹⁴For instance, Topel and Ward (1992) find that after ten years American workers are on their 7th job on average, while German workers have worked for only 3 employers (Dustmann and Meghir (2003)). See also table 1.1.

distinction between job-to-job and job-to-unemployment mobility is important. When moving from job-to-job workers gain search capital, whereas they lose search capital when becoming unemployed. I first derive the restrictions the search model imposes on wage gains of stayers, job-to-job and job-to-unemployment movers as well as on the covariance structure of wages and wage growth. I then propose a simple method of decomposing the variance of the log-wage residual into the variance of ability, match quality and a transitory component.

Our main findings can be summarised as follows. Mobility is substantially lower in Germany than in the US for all education groups. German workers are particularly less likely to become unemployed. In Germany, apprentices less mobile than unskilled workers. Yet, there are important similarities between US and German workers. Match-specific productivity plays an important role in both countries for all education groups. Furthermore, the proportion of the variance that can be attributed to match declines with experience, while the proportion that can be attributed to ability increases with experience in both countries for all education groups. There is also some evidence for learning about match quality in both countries. Heterogeneity in the rate of firm-specific human capital accumulation, in contrast, is not an important feature of the data in either country, as the autocovariance of the within-job wage growth residual is close to zero and uniformly negative at lag greater than 1.

American and German workers mostly differ with respect to the variance of ability and transitory wage component. The variance of ability is considerably higher in US for all education groups. Furthermore, learning about ability plays an important role in the US, but only a modest role in Germany. The proportion of the variance attributable to match quality is considerably higher in Germany than in the US. This may indicate that search frictions play a more important role in Germany. As expected, the variance of the transitory wage component is substantially lower in the German administrative data than in the US survey data due to measurement error.

There are also interesting differences between unskilled workers and apprentices in Germany. First, the variance of the quality of the match is almost three times as high for unskilled workers than for apprentices at low levels of experience. This difference all

but disappears with experience. This is consistent with the idea that job search plays a more important role for unskilled workers than for apprentices. Second, the variance of the within-job wage growth residual is considerably higher for unskilled workers than for apprentices at low tenure levels, particularly so at the first job. This difference all but disappears with tenure. One reason for the lower mobility of apprentices than for unskilled workers may thus be a higher precision of the quality of the match.

Chapter 5 concludes the thesis with a review of the key findings and some suggestions for future research.

Chapter 2

Testing for Asymmetric Employer Learning

In a world where the productivity of labour market participants is difficult to observe, their individual career path and employment history forms a significant basis for any company's hiring decisions. The longer a job-candidate's employment history, the more information a recruiting firm can draw upon in assessing the candidate's ability to do the job she is being considered for. Some information about the candidate's ability, such as her education or her CV, is available to all recruiting firms. Other information may be available only to the candidate's current employer. As an illustration, consider a lawyer. Clearly, the incumbent firm will use the number of cases the lawyer wins and loses to update its belief about her productivity. This information may also be available to outside firms. The incumbent firm is likely to be better informed about the lawyer's ability to work in a team, her ability to cope with stress, etc. Recent research by Farber and Gibbons (1996) and Altonji and Pierret (2001) assumes that recruiting firms have the same information about workers' productivity as their current employers. In the literature, this assumption is termed "symmetric employer learning". In contrast, earlier work by Gibbons and Katz (1991) assumes that recruiting firms have an informational disadvantage relative to workers' current employers. This assumption is termed "asymmetric employer learning". Both learning hypotheses have found empirical support.

The question of this paper is: Is employer learning symmetric or asymmetric? We develop a simple model, and compare the empirical implications of two polar cases: the case when all learning is public (symmetric learning), and the case when all learning is private (asymmetric learning). We show that existing tests that have been derived under one learning hypothesis are also consistent with the other, and propose new tests that allow to discriminate between the two hypotheses. A key difference between our and most other learning models is endogenous mobility: In our model, there are endogenous job-to-job as well as job-to-unemployment transitions¹. This paper therefore also adds to the literature on worker mobility by analyzing how workers' general productivity as well as learning about this productivity affects the probability of a job-to-job and job-to-unemployment transition². The way we endogenize mobility is

¹The symmetric learning models by Farber and Gibbons (1996) and Altonji and Pierret (2001) do not consider worker mobility. Similarly, in many asymmetric learning models, such as Waldman (1984) and (1996), and Bernhardt (1995), no worker switches firms. Other asymmetric learning models, such as Greenwald (1986) and Gibbons and Katz (1991), assume that some workers leave the firm for exogenous reasons. Bernhardt and Scones (1993) is the only asymmetric learning model that allows for match-specific productivity.

To the best of our knowledge, no model with learning about general productivity as well as endogenous job-to-job and job-to-unemployment transitions exists. Montgomery (1999), Kugler and Saint-Paul (2000) and Berkovitch (1990) are examples for asymmetric learning models in which the market regards unemployment, or unemployment duration, as a bad signal. In these models there are either no job-to-job transitions, as in Montgomery (1999), or job-to-job transitions are exogenous, as in Kugler and Saint-Paul (2000).

²There are few theoretical contributions that analyze how ability -as opposed to match-specific productivity- affects mobility. Neal (1998) proposes a model in which more able workers optimally choose more specialized jobs. His model predicts a negative correlation between ability and mobility. Neal (1998) does not analyse how learning about ability as well as the type of learning affects mobility, neither does he distinguish between job-to-job and job-to-unemployment transitions.

The empirical literature has long recognized the importance of ability for mobility decisions. For instance, the literature on the returns to tenure (e.g. Topel (1991), Altonji and Shakotko (1987), Dustmann and Meghir (2003)) is concerned with an "ability bias": OLS estimates may overestimate true returns to tenure because more able workers stay with the same employer longer. There is also an ongoing discussion on whether a declining job hazard rate is driven by worker heterogeneity, or reflects true negative duration dependence. See Farber (1999) for a summary.

similar to Neal (1998) and Acemoglu and Pischke (1998). We assume that workers' utility depends not only on the unemployment benefit and the wage they earn, but also on non-pecuniary job characteristics, such as the distance to work, task variety, personal relationships with co-workers, etc.

Distinguishing between symmetric and asymmetric learning is important for several reasons. Firstly, we show that under asymmetric learning the labour market fails to efficiently allocate workers to firms. As the incumbent firm has superior information about workers' ability, wage formation in incumbent firms will differ from wage formation in outside firms. The job that yields the highest utility to the worker is thus not necessarily the job for which the worker is best suited³.

Secondly, the type of employer learning also has implications for the wage structure. Under asymmetric learning the informational advantage allows incumbent firms to pay wages below productivity so that general productivity is rewarded as if it were match- or firm-specific. This has important consequences for educational policies. As employers earn a higher rent on more able workers, *firms* have an incentive to pay for general training if training and ability are complements (Acemoglu and Pischke (1998)). On the other hand, asymmetric learning distorts the incentive for *workers* to pay for skill acquisition.

Thirdly, easily observable signals about workers' productivity, such as workers' education, will have an impact on wage determination much longer under asymmetric than under symmetric learning. There is thus more scope for statistical discrimination under asymmetric learning. This may have consequences for the effectiveness of active labour market programs. If employer learning is asymmetric and recruiting employers continue to use easily observable signals to statistically discriminate among workers, program participation may be used as a signal for workers' productivity. If program participants are less able on average than non-participants, they may become stigmatized. Active labour market programs will be less effective and may even harm workers.

³This is related to Waldman's (1984) and Bernhardt's (1995) result that under asymmetric learning the promotion rule, and thus job assignment within firms, is inefficient.

Our testing strategy proceeds in two steps. Both tests make use of a variable that is an indicator for workers' ability, but not observed by employers, such as a test score. A well known consequence of asymmetric learning is that workers who switch firms are of lower ability than workers who stay with their employer (e.g. Greenwald (1986), Gibbons and Katz (1991)). Our first test is a test for a lower ability of movers. We show that a distinction between job-to-job and job-to-unemployment movers is crucial here. Both symmetric and asymmetric learning predict a lower ability of job-to-unemployment movers than of stayers. The two learning hypotheses differ with respect to job-to-job movers. Symmetric learning implies that job-to-job movers are as able as stayers, while asymmetric learning implies that job-to-job movers are less productive than stayers. We test these implications directly by comparing the average test score of the three groups of workers.

Our second test is a modification of Farber and Gibbons' (1996) and Altonji and Pierret's (2001) test for symmetric learning. We show that symmetric and asymmetric learning differ with respect to how a hard- and easy-to-observe variable, such as ability and education, affect wages for those who stay and move. Under symmetric learning incumbent and outside firms have the same information about workers' ability and education. Hence, the test score and education affect wages for movers in the same way as wages for stayers. In contrast, under asymmetric learning incumbent firms use the newly acquired information about workers' ability, while outside firms continue to use education as a signal for workers' ability when determining wages. The coefficient on the test score is thus higher for stayers than for movers, while the coefficient on education is higher for movers than for stayers.

Our empirical analysis is based on the same data set that has been used by Farber and Gibbons (1996) and Altonji and Pierret (2001) to test for symmetric learning, the NLSY79. The data set contains the complete work history, including all job-to-job and job-to-unemployment transitions, of a cohort of workers from labour market entry onwards. It thus focuses on young workers for whom learning and mobility should matter most. We follow Farber and Gibbons (1996) and Altonji and Pierret (2001) and use the Armed Forces Qualification Test (AFQT) as a variable that is correlated with

ability but not observed by employers. The AFQT- score provides a summary measure for basic literacy and numeric skills and is thus an indicator for workers' cognitive ability. Our tests are therefore best understood as tests for employer learning about workers' cognitive ability as opposed to about workers' non-cognitive ability, such as communication skills, etc.

We perform all tests separately by workers' education. We do this because our model implies that asymmetric learning has stronger consequences for workers with more education if output of better educated workers is more sensitive to ability. We find some empirical support for this assumption. Moreover, a distinction between education groups turns out to be very important empirically. For high school dropouts and graduates we find no evidence for asymmetric learning. For university graduates, in contrast, the empirical evidence is supportive of asymmetric learning. Our results thus indicate that the way how the market acquires new information about workers' productivity differs across education groups.

The plan of the chapter is as follows. Section 2.1 reviews existing tests for symmetric and asymmetric employer learning. Section 2.2 develops the learning model. Section 2.3 contrasts the empirical implications of symmetric learning with those of asymmetric learning. Section 2.4 describes the data. Section 2.5 reports results, and discusses whether the results are also compatible with explanations other than asymmetric learning. Section 2.6 concludes.

2.1 Existing tests for employer learning

In this section we briefly describe existing tests for symmetric and asymmetric learning. Gibbons and Katz (1991) were the first to provide a test for asymmetric learning. Others, such as Grund (1999) and Doiron (1995), have used their methodology to test for asymmetric learning in other countries.

Gibbons and Katz's test for asymmetric learning is based on the idea that under asymmetric learning the market will regard a layoff as a bad signal. No such stigma is attached to workers who are displaced from their firm for exogenous reasons. This

implies a lower post-displacement wage for laid-off workers than for exogenous movers. Exogenous movers and laid-off workers look the same when they enter the labour market. Hence, asymmetric learning additionally implies the same pre-displacement wage for both groups of workers. Taken together, asymmetric learning predicts a greater wage loss for laid-off workers than for exogenous movers. Using data from the CPS Displacement Survey, Gibbons and Katz (1991) find strong support for these hypotheses⁴.

As Gibbons and Katz' model, our model is a two period model. An important difference is that there are no layoffs in our model. Instead, we distinguish between job-to-job and job-to-unemployment movers. We show that under symmetric learning job-to-unemployment movers are less able and earn a lower post-displacement wage than exogenous movers. Hence, if laid-off workers and job-to-unemployment workers are correlated, then symmetric learning leads to the same implications as Gibbons and Katz's asymmetric learning model.

Farber and Gibbons (1996) were the first to derive empirical implications of symmetric learning. They were subsequently extended by Altonji and Pierret (2001). Bauer and Haisken-DeNew (2001) and Galindo-Rueda (2002) use a similar methodology to test for employer learning in Germany and the UK, respectively⁵. Learning models predict that at labour market entry when firms have only imprecise information about workers' productivity, firms rely on easy-to-observe variables that are correlated with productivity, such as education, when determining wages. As employers acquire new information about workers' productivity, wages become increasingly dependent on productivity. Hence, the coefficient on a variable correlated with productivity that is unobserved by employers, but observed by the econometrician, should increase with experience. By the same argument, the coefficient on an easy-to-observe variable which is correlated with ability should decrease with experience if the hard to observe variable

⁴These results have recently been questioned by Krashinsky (2002). Using data from the NLSY79, he shows that accounting for establishment size removes virtually all of the difference in the wage losses for the two groups of workers.

⁵Both authors also test for asymmetric employer learning.

is included (Altonji and Pierret (2001)), and remain constant if the hard-to-observe variable is not included in the wage regression (Farber and Gibbons (1996)). Farber and Gibbons derive the additional implication that wages should follow a martingale. Both Farber and Gibbons and Altonji and Pierret find empirical support for symmetric learning.

Our model is simpler than Farber and Gibbons and Altonji and Pierret's model because we consider two periods only. On the other hand, it is richer because of mobility. This allows to analyse how hard- and easy-to-observe variables differently affect wages for movers and stayers under the two learning hypotheses. We show that the implications of symmetric learning are also compatible with asymmetric learning, and modify them such that it is possible to distinguish between the two learning hypotheses.

2.2 The learning model

This section sets up a simple learning model and solves it under the assumption of symmetric and asymmetric learning. Many assumptions are discussed in section 2.3.4.

2.2.1 Description of environment

There are two periods. There is a continuum of workers. Each period two firms compete for workers. This assumption has been made for simplicity only and can be generalized to N firms. Workers maximise expected utility, firms maximise expected profits. There is no discounting. Workers (and firms) are risk-neutral.

Productivity Workers differ with respect to 'ability' and 'education'. Workers' education is known to all employers, and thus an 'easy-to-observe' variable. Ability may in principle comprise all those characteristics workers bring to the labour market which are hard to observe by employers, such as cognitive ability, communication skills, etc. For simplicity, workers are either educated or uneducated. Educated and uneducated workers are either of low or high ability. We denote education by the superscript

k , $k = u, e$, and ability by the subscript i , $i = L, H$. Table 2.1 specifies workers' productivity y_i^k .

Table 2.1: Specification of workers' productivity

	uneducated	educated
low ability	$y_L^u = a_L$	$y_L^e = a_L$
high ability	$y_H^u = a_H$	$y_H^e = a_H + s$
proportion with low ability	p^u	p^e

We make the normalizing assumption that education increases the productivity only of high, but not of low ability workers. What matters is that workers of high ability benefit more from acquiring education than workers of low ability, i.e. $s > 0$ (*single crossing property*). Later, we propose a simple test for the single crossing property. Recent research by Carneiro et al. (2003) is supportive of such a complementary between ability and education. He finds that there is considerable heterogeneity in the returns to education as well as a positive correlation between the returns to education and the fixed effect in a wage regression.

We further assume that the proportion of low ability workers is lower among workers with education, i.e. $p^e < p^u$. This assumption implies that education is informative about ability. There is ample empirical evidence for a positive correlation between education and ability. For Sweden, Meghir and Palme (2001) find that various measures of a child's cognitive ability subsequently increase educational attainment. For the US, Heckman and Vytlačil (2001) as well as Cawley et al. (2001) document a strong correlation between ability measures and schooling.

We also assume that workers neither accumulate general nor firm-specific human capital.

Information structure In the first period information is symmetric with respect to workers and firms, i.e. workers do not have superior information to firms about

their ability⁶. In the second period either all firms (symmetric learning) or only the incumbent firm (asymmetric learning) get to know workers' ability. The proportion of workers with low and high ability among educated and uneducated workers is common knowledge.

Unemployment Each period workers are either unemployed or work. When unemployed, workers receive an unemployment benefit b . Alternatively, b may be thought of as disutility from work. For simplicity, we assume that the unemployment benefit is the same for educated and uneducated as well as for low and high ability workers. Our results continue to hold if the difference between the utility from unemployment is smaller than the difference between the productivity of high and low ability workers, i.e. $b_H^k - b_L^k < y_H^k - y_L^k$, $k = u, e$. This condition implies that high ability workers have a comparative advantage at work. It is satisfied in a social security system that guarantees a minimum level of income to the unemployed.

Job-to-job and job-to-unemployment mobility A crucial feature of our model is that workers do not only derive utility from pecuniary characteristics, such as wages or unemployment benefits, but also from non-pecuniary characteristics. Only the worker, but not the firm, observes the utility from non-pecuniary characteristics. We specify workers' utility from working at firm j as a linear function of the wage, w_j , and the non-pecuniary characteristic, m_j :

$$U(\text{work at firm } j) = w_j + m_j, \quad j = 1, 2.$$

The non-pecuniary characteristic may reflect distance to work, personal relationships with co-workers and superiors, health care programs etc. Similarly, workers' utility

⁶Even with symmetric information, more able workers may be more likely to attend school. For instance, suppose that workers as well as firms receive a signal about workers' ability, such as workers' school grades. Since ability is higher valued at skilled jobs, workers with a positive signal will be more likely to go to university. Hence, the proportion of low ability workers will be higher among uneducated workers than among educated workers. Nevertheless, school attendance does not provide additional information about workers' ability to firms.

from unemployment is a linear function of the unemployment benefit b and the non-pecuniary characteristic m_0 :

$$U(\text{unemployment}) = b + m_0.$$

Here, the non-pecuniary characteristic may reflect personal taste for working, or non-labour income, etc. We assume that the three draws of non-pecuniary characteristics (i.e. for the incumbent firm, the outside firm and unemployment) are independent draws from a known distribution. Let \tilde{g} and \tilde{G} denote the probability density function and the cumulative distribution function of this distribution. It is independent of workers' skill and ability. We show our theoretical results under the assumption that \tilde{G} follows an extreme value distribution, i.e. $\tilde{G}(m) = \exp\{-e^{-\frac{m}{\lambda}}\}$. This distributional assumption allows to derive particularly simple expressions for the probability of staying, moving from job-to-job and becoming unemployed. We further assume that each period workers draw a new measure of non-pecuniary job characteristics at each job. Hence, non-pecuniary job characteristics are transitory.

Wage offers We assume that each period firms compete for workers by simultaneously making wage offers to workers⁷. We therefore rule out long-term contracts to workers. This is in line with our assumption that workers are risk-neutral. Consequently, workers do not benefit from long-term contracts which insure them against low realizations of their ability, such as in Harris and Holmstrom (1982).

Timing The exact timing of events is as follows.

1. At the beginning of period 1 two firms simultaneously make a wage offer to the worker. Firms observe workers' education, but not their ability.

⁷Instead of two firms simultaneously making wage offers to the worker, we can alternatively assume that first the outside firm makes a wage offer. The incumbent firm observes the worker's outside offer, and then makes a counter-offer. This is the standard assumption about wage determination in the asymmetric learning literature, as e.g. in Gibbons and Katz (1992). It is important, however, that the outside firm does not observe the incumbent firms' wage offer.

2. Workers observe their utility from non-pecuniary characteristics at each firm as well as from unemployment. They choose the job offer that yields the highest utility, and unemployment if the utility from unemployment exceeds the utility from the preferred job offer.
3. Production takes place. Either all firms (symmetric learning) or only the incumbent firm (asymmetric learning) learn about workers' ability.
4. At the beginning of period 2 the incumbent firm and one outside firm simultaneously make a wage offer to the worker.
5. Workers draw a new measure of non-pecuniary job characteristics for both the incumbent firm and the outside firm and unemployment. Workers choose the job offer that yields the highest utility, and unemployment if the utility from unemployment exceeds the utility from the preferred job offer.
6. Production takes place. At the end of the second period workers retire.

Section 2.2.1 solves the model under the assumption of symmetric learning, section 2.2.3 under the assumption of asymmetric learning. Symmetric learning

2.2.2 Symmetric learning

Workers' quit decision

We solve the model backwards and begin with the worker's optimal quit decision in the second period. A worker may either stay with his current firm, move to the outside firm (job-to-job transition) or become unemployed (job-to-unemployment transition). Consider a worker with given education and ability level. To simplify the notation, we drop the superscript k and subscript i . Let w denote the wage offer of the incumbent firm and v the wage offer of the outside firm. The utility from non-pecuniary characteristics at the incumbent and outside firm as well as when unemployed is denoted by m_1 , m_2 and m_0 , respectively. A worker becomes unemployed if the utility from unemployment

exceeds the utility from either job, i.e. if $b + m_0 > w + m_1$ and $b + m_0 > v + m_2$. For the extreme value distribution, this probability equals

$$\Pr(\text{job-to-unemployment}) = \frac{e^{b/\lambda}}{e^{b/\lambda} + e^{w/\lambda} + e^{v/\lambda}}.$$

By the same argument, the worker stays with the employer if $w + m_1 > v + m_2$ and $w + m_1 > b + m_0$. The probability of staying can be computed as

$$\Pr(\text{stay}) = \frac{e^{w/\lambda}}{e^{b/\lambda} + e^{w/\lambda} + e^{v/\lambda}}.$$

The probability of a job-to-job transition can be similarly derived as

$$\Pr(\text{job-to-job}) = \frac{e^{v/\lambda}}{e^{b/\lambda} + e^{w/\lambda} + e^{v/\lambda}}.$$

Wage determination in the second period

We next analyse wage determination in the second period. Under symmetric learning incumbent as well as outside firms observe both workers' ability and education. Hence, wages depend on ability as well as on education. Firms set wages by maximizing expected profits. Expected profits depend on the probability that a firm attracts a worker and the rent a firm earns per retained worker. Firms face the trade-off between a higher chance of attracting a worker and a lower profit per worker⁸. The incumbent firm maximises

$$\max_{w_i^k} \frac{e^{w_i^k/\lambda}}{e^{b/\lambda} + e^{w_i^k/\lambda} + e^{v_i^k/\lambda}} (y_i^k - w_i^k), \quad i = L, H; k = u, e, \quad (2.1)$$

while the outside firm maximises

$$\max_{v_i^k} \frac{e^{v_i^k/\lambda}}{e^{b/\lambda} + e^{w_i^k/\lambda} + e^{v_i^k/\lambda}} (y_i^k - v_i^k), \quad i = L, H; k = u, e. \quad (2.2)$$

In appendix A.1 we show that an equilibrium exists and is unique, and that first order conditions are sufficient for a maximum. Observe that both firms solve the same maximization problem. Hence, incumbent and outside firms offer the same wage in equilibrium, i.e. $v_i^k = w_i^k$. As workers stay with a positive probability with the incumbent firm even if they receive a higher outside wage offer, firms have monopsonic power

⁸The bidding process is similar to Fudenberg and Tirole (2000).

and pay wages below productivity. Since incumbent and outside firms offer the same wage, workers move from job-to-job only if the utility from non-pecuniary characteristics at the incumbent firm, m_1 , exceeds the utility from non-pecuniary characteristics at the outside firm, m_2 . Hence, under symmetric learning job-to-job mobility is efficient.

Wage determination in the first period

We analyse wage determination in the first period in appendix A.3. At this point, it is important to bear in mind that in the first period firms only observe workers' education, but not their ability. Consequently, wages in the first period depend on education only, but not on ability.

2.2.3 Asymmetric learning

This section analyses the model under the assumption of asymmetric learning.

Wage determination in the second period

We start with wage determination in the second period. We analyse wage determination in the first period in appendix B.4. Under asymmetric learning the incumbent firm observes the worker's ability. It uses this information when determining wages and offers a higher wage to more able workers. The outside firm, on the other hand, is only informed about workers' education, but not ability. It thus offers the same wage for high and low ability workers, but a different wage for educated and uneducated workers. The incumbent firm maximises

$$\max_{w_i^k} \frac{e^{w_i^k/\lambda}}{e^{b/\lambda} + e^{w_i^k/\lambda} + e^{v^k/\lambda}} (y_i^k - w_i^k), \quad i = L, H; k = u, e. \quad (2.3)$$

Since low ability workers receive the same outside wage offer, but a lower wage offer from the incumbent firm, they are more likely to switch jobs than high ability workers. Outside firms have to take this into account when determining which wage to offer.

Outside firms maximise

$$\begin{aligned} \max_{v^k} \quad & p^k \frac{e^{v^k/\lambda}}{e^{b/\lambda} + e^{w_L^k/\lambda} + e^{v^k/\lambda}} (y_L^k - v^k) + \\ & (1 - p^k) \frac{e^{v^k/\lambda}}{e^{b/\lambda} + e^{w_H^k/\lambda} + e^{v^k/\lambda}} (y_H^k - v^k), \quad k = u, e. \end{aligned} \quad (2.4)$$

In appendix B.1 we show that an equilibrium exists and is unique, and that first order conditions are sufficient for a maximum. In equilibrium we have $w_L^k < v_i^k < w_H^k$: Outside firms offer a wage in between the wage offer of incumbent firms for low and high ability workers. Since outside and incumbent firms offer a different wage to the same worker, asymmetric learning results in an inefficient allocation of workers to jobs. Low ability workers are too mobile compared to the efficient level since $w_L^k - v^k < 0$. This wage difference can be interpreted as a staying cost. High ability workers, on the other hand, are less mobile than optimal. The wage difference $w_H^k - v^k$ can be regarded as a mobility cost.

An important difference between our and existing asymmetric learning models is that in our model the wage offer of the incumbent firm depends on workers' ability. This difference arises because of non-pecuniary job characteristics⁹: In the presence of non-pecuniary characteristics firms can no longer guarantee to retain a worker simply by matching her outside offer. As more able workers are more valuable to the firm, it becomes profit-maximizing for the incumbent firm to pay a higher wage to more able workers. Non-pecuniary characteristics also imply that it is not only low ability workers who switch firms.

2.3 Empirical implications

This section compares the empirical implications of both learning hypotheses. Suppose the data set contains a test score that is correlated with ability. This test score is not observed by employers. For simplicity, we assume that the test score can take

⁹Note that in Acemoglu and Pischke (1998) workers also care about non-pecuniary job characteristics. Acemoglu and Pischke (1998), however, assume that incumbent firms have to offer the same wage to low and high ability workers.

two values, low and high. Let q_L^{LT} denote the probability that a worker with low ability scores low in the test. Similarly, q_H^{HT} is the probability that a high ability worker achieves a high test score. We assume that the signal about workers' ability is symmetric, i.e. $q_L^{LT} = q_H^{HT}$. For the test score to be informative about ability, we require $q_L^{LT} > 0.5$.

We first contrast symmetric and asymmetric learning with respect to the average ability - and thus test score- of three groups of workers: stayers, job-to-job and job-to-unemployment movers (section 2.3.1). We then modify Farber and Gibbons (1996) and Altonji and Pierret's (2001) implications of symmetric learning such that it is possible to distinguish between the two learning hypotheses (section 2.3.2), and propose a simple test for the single crossing property (section 2.3.3). Finally, we summarize our results and discuss the importance of our assumptions (section 2.3.4).

2.3.1 Stayers, job-to-job and job-to-unemployment movers

Symmetric learning

We start with a comparison of the average ability of stayers, job-to-job and job-to-unemployment movers under symmetric employer learning.

Job-to-job movers First, consider job-to-job movers. Under symmetric learning incumbent firms offer the same wage as outside firms, i.e. $w_i^k = v_i^k$. Hence, the probability of staying is the same as the probability of moving from job-to-job, for both low and high ability workers. Consequently, on average stayers are as able and earn the same wage as job-to-job movers. This implies that the test score of stayers and job-to-job movers is the same on average.

Job-to-unemployment movers Next, consider job-to-unemployment movers. Symmetric learning implies a lower average ability -and thus a lower average test score- for job-to-unemployment movers than for stayers and job-to-job movers. This is because high ability workers have a higher utility from working than low ability workers, but on average the same utility from unemployment. Low ability workers are thus more

likely to become unemployed. Hence, on average job-to-unemployment movers are less able -and have a lower test score- than stayers. This result continues to hold if high ability workers have a comparative advantage at work, i.e. if $y_H^k - y_L^k > b_H^k - b_L^k$. We summarize our results in the following proposition.

Proposition 1 *Under symmetric learning (i) job-to-job movers have the same test score and earn the same wage on average as stayers. (ii) Job-to-unemployment movers have lower test scores on average than stayers and job-to-job movers.*

Proof. See appendix A.2. ■

Asymmetric learning

We next compare the average ability of stayers, job-to-job and job-to-unemployment movers under asymmetric learning.

Job-to-job movers First, consider job-to-job movers. Under asymmetric learning high ability workers receive the same outside offer, but a higher wage offer from the incumbent firm than low ability workers. Clearly, low ability workers are more likely to move from job-to-job. Stayers are thus more able on average - and have a higher test score- than job-to-job movers. Furthermore, the difference is higher for educated workers if productivity of educated workers is more sensitive to ability (single crossing property). The reason is that an increase in the productivity of high ability workers reduces the job-to-job transition rate of high ability workers, thereby rising the productivity difference between job-to-job movers and stayers. The single crossing property thus implies a stronger negative selection of job-to-job movers for educated workers.

Job-to-unemployment movers Next, consider job-to-unemployment movers. As symmetric learning, asymmetric learning implies that job-to-unemployment movers are less able -and thus have a lower test score- than stayers. This is because low and high ability workers have the same utility from unemployment, but high ability workers have a higher utility from staying than low ability workers. We summarize

Proposition 2 *Under asymmetric learning (i) stayers have higher test scores on average than job-to-job movers. The difference is higher for educated workers. (ii) Job-to-unemployment movers have lower test scores on average than stayers.*

Proof. See appendix B.2. ■

Note that asymmetric learning may in fact result in a lower ability of stayers than of job-to-unemployment movers if the utility from unemployment is higher for more able workers, i.e. $b_H > b_L$, as high and low ability workers receive the same outside wage. Symmetric and asymmetric learning thus differ with respect to job-to-job movers, and not, as Gibbons and Katz's (1991) test for asymmetric learning may suggest, with respect to job-to-unemployment movers.

2.3.2 Modification of existing tests for symmetric learning: Learning by experience or tenure?

We next show how Farber and Gibbons (1996) and Altonji and Pierret's (2001) implications of symmetric learning can be modified such that it is possible to distinguish symmetric from asymmetric learning. Symmetric learning implies that in a wage regression the coefficient on the test score increases with experience (Farber and Gibbons (1996)). The coefficient on workers' education, on the other hand, decreases with experience if the test score is included in the wage regression (Altonji and Pierret (2001)). We show that these implications are also consistent with asymmetric learning. Symmetric learning differs from asymmetric learning with respect to how education and ability affect wages for those who stay and those who move.

Under symmetric learning incumbent and outside firms have the same information about workers' ability and education. Hence, ability and education have the same impact on the wage offer of incumbent and outside firms. The coefficient on the test score and on education should therefore be the same for stayers and movers.

Under asymmetric learning, in contrast, ability and education differently affect the wage offer of incumbent and outside firms. This is a direct consequence of informational asymmetries: As only incumbent firms learn about workers' ability, only the wage offer

of incumbent firms, but not that of outside firms, reflect the new information about workers' ability. Hence, the coefficient on the test score is higher for those who stay than for those who move. This implies a positive coefficient on the interaction between tenure and the test score¹⁰. The coefficient on education, on the other hand, should be higher for movers than for stayers if the test score is included in the wage regression. This is simply the other side of the same coin: Outside firms continue to use workers' education as a signal for workers' ability while incumbent firms offer a wage that correctly reflects workers' ability. This implies a negative coefficient on the interaction between tenure and education.

We summarize our results in the following proposition.

Proposition 3 *Under symmetric learning the coefficient on the interaction between tenure and the test score as well as tenure and education are zero. Under asymmetric learning (i) the coefficient on the interaction between tenure and the test score is positive. (ii) If the test score is included in the wage regression, the coefficient on the interaction between tenure and education is negative.*

Proof. See appendix B.2. ■

Note that the assumption that employers do not observe the test score is crucial here. If they did, they would use this information to update their belief about workers' ability, and wages for both movers and stayers would reflect this information. More precisely, what is required is that the test score contains some information about workers' ability that is not available to employers. This is not to say that the econometrician has more information about workers' ability than employers.

Also note that Farber and Gibbons (1996) and Altonji and Pierret's (2001) tests symmetric employer learning are consistent with asymmetric employer learning. Under pure symmetric learning, all learning takes place by *experience*. In contrast, under pure asymmetric learning, all learning takes place by *tenure*. Due to the correlation between

¹⁰Bauer and Haisken-DeNew (2001) also use the interaction between tenure and fathers' education (as a variable that is correlated with ability but unobserved by employers) to test for asymmetric learning. However, they do not provide a theoretical justification for this approach.

tenure and experience, asymmetric learning implies a positive interaction between the test score and experience if tenure is not controlled for. Existing tests for symmetric learning are thus best understood as tests for employer learning in general.

2.3.3 Test for the single crossing assumption

We next provide a simple test for the assumption that the productivity of educated workers is more sensitive to ability than the productivity of uneducated workers when learning is symmetric and asymmetric, respectively. Under symmetric learning wages in the second period, but not wages in the first period, reflect workers' ability, independently of whether the worker stays or moves. Hence, if the productivity of educated workers is more sensitive to ability, then the increase of the test score with *experience* should be higher for educated workers. This implies a positive coefficient on the interaction between education, the test score, and *experience*. Under asymmetric learning, in contrast, only wages for stayers reflect workers' ability. Hence, if the productivity of educated workers is more sensitive to ability, then the increase of the test score with *tenure* should be higher for educated workers. This implies a positive coefficient on the interaction between education, the test score, and *tenure*. We summarize

Proposition 4 *(i) Under symmetric learning the coefficient on the interaction between experience, education, and the test score is positive. (ii) Under asymmetric learning the coefficient on the interaction between experience, education, and the test score is positive.*

Proof. See appendix B.3. ■

2.3.4 Summary and discussion of assumptions

Table ?? summarizes the implications of the pure symmetric and asymmetric learning models. First, symmetric and asymmetric learning differ with respect to job-to-job movers and not -as existing tests for asymmetric learning may suggest- with respect to job-to-unemployment movers. Under symmetric learning stayers are as able on

average as job-to-job movers, whereas under asymmetric learning stayers are more able on average than job-to-job movers. Second, symmetric learning differs from asymmetric learning with respect to how an easy- and a hard-to-observe variable that is correlated with ability, such as education and a test score, affect wages for stayers and movers. Under symmetric learning wages for movers and stayers are determined in the same manner. Hence, the coefficient on the test score and on education is the same for movers and stayers, and the coefficients on the interactions between the test score and tenure and between education and tenure, respectively, should be zero. In contrast, under asymmetric learning incumbent firms observe workers' ability and offer a wage that depends on ability. Outside firms, on the other hand, continue to rely on education when determining wages. Hence, the coefficient on the test score should be higher for stayers whereas the coefficient on education should be higher for movers. This implies a positive coefficient on the test score-tenure interaction and a negative coefficient on the education-tenure interaction.

Table 2.2: Empirical implications of symmetric and asymmetric learning

Symmetric Learning	Asymmetric Learning
-job-to-job movers as able as stayers	- job-to-job movers less able than stayers → wage job-to-job movers < wage stayers → difference higher for educated workers
-job-to-unemployment movers less able than stayers	-job-to-unemployment movers less able than stayers
- test · tenure = 0	- test · tenure ≥ 0
- education · tenure = 0	- education · tenure ≥ 0
- test · education · tenure = 0	- test · education · tenure ≥ 0

So far, we have only discussed the two polar cases of pure symmetric and asymmetric learning. In reality, learning may be both symmetric and asymmetric. In this case the coefficients on both the experience-AFQT- score and tenure-AFQT- score-interaction should be positive, while the coefficients on both the experience-education and tenure-education interaction should be negative. Our test therefore is a test whether learning

partly asymmetric.

We next discuss the importance of our assumptions for these results.

Human capital accumulation We have assumed that workers accumulate neither general nor firm-specific human capital. As general human capital does not affect workers' quit decision (under symmetric learning), our tests for asymmetric employer learning are not affected by general human capital accumulation. A model in which more able workers accumulate more firm-specific human capital, however, is consistent with many implications of asymmetric learning. We discuss this in detail in section 2.5.3.

Symmetric information in the first period We have assumed that in the first period information is symmetric with respect to workers and firms. This assumption affects wage determination in the first period, but not firms' and workers' behavior in the second period. The empirical implications of asymmetric learning are therefore unchanged. In appendix A.3, we show that this assumption implies that the probability of being unemployed in the first period is the same for low and high ability workers. Unemployment in the first period is therefore not informative about workers' ability.

Next, suppose that in the first period workers know their type, but employers do not. Suppose further that the utility from unemployment is the same for low and high ability workers. High ability workers will earn a higher wage than low ability workers in the second period. Accepting a job offer is thus more attractive for high than for low ability workers. Consequently, unemployment in the first period will provide useful information for firms about workers' ability. This -of course- will affect workers' decision whether to accept a job or remain unemployed, and firms' decision which wage to offer in the first period. A detailed analysis of these issues is beyond the scope of this paper¹¹. The maximisation problem in the second period, however, is the same.

¹¹Rodriguez-Planas (1998) considers a model with asymmetric information in which it is high ability workers who remain unemployed longer. The idea is that accepting a low wage signals a low ability to the market. In our model high ability workers may be more likely to remain unemployed if their utility from unemployment is higher than that of low ability

Match-specific productivity In our model, workers switch jobs because their *utility* differs across firms. Instead, we could have assumed that workers' *productivity* differs across firms. This clearly complicates the firm's maximization problem. We believe, however, that match-specific productivity has little impact on the crucial implications of asymmetric learning, i.e. a lower ability of job-to-job movers and a higher coefficient on the test score for stayers than for movers. These implications follow from the informational asymmetry, i.e. incumbent firms offer a higher wage to more able workers, but outside firms don't. Incumbent firms have an incentive to do so even with match-specific productivity.

We would also like to note that a higher wage growth for job-to-job movers, than for stayers -as observed in many studies, such as Topel and Ward (1992)- should not be seen as evidence against asymmetric learning. Even if workers on average gain from job mobility, it may be the less able workers who are more likely to leave.

Extreme value distribution We have assumed that non-pecuniary job characteristics are drawn from an extreme value distribution. We have done so because it allows to derive particularly simple expressions for the probability of staying, moving from job-to-job and moving from job-to-unemployment. Since incumbent firms offer a higher wage for more able workers for a wide variety of distributions of non-pecuniary characteristics, we believe that this assumption is not essential.

Transitory non-pecuniary job characteristics In our model, non-pecuniary job characteristics are transitory. This implies that the distribution from which non-pecuniary characteristics are drawn is the same for incumbent and outside firms. It is because of this assumption that under symmetric learning incumbent and outside firms offer the same wage.

We have simulated the model under the assumption of permanent non-pecuniary job characteristics. We found that -opposite to the predictions of the asymmetric learning model- job-to-job movers are more able than stayers under symmetric learning. However, at this point we cannot rule out that symmetric learning is consistent with a

workers.

lower ability of job-to-job movers than of stayers when non-pecuniary job characteristics are permanent. A lower ability of job-to-job movers may therefore best be regarded as a necessary implication of asymmetric learning, which may not be sufficient to distinguish symmetric from asymmetric learning¹².

2.4 Data description and variables

The data comes from the 1979 National Longitudinal Survey of Youth (NLSY79). At the time we started this analysis the data was available from 1979 to 1996. The data is well suited for our purposes. First, it contains the complete work history, including all job-to-job and job-to-unemployment transitions, of a cohort of workers from labour market entry onwards. It thus focuses on young workers for whom learning and worker mobility should matter most. Second, for each respondent it contains a test score variable that is unlikely to be observed by employers, which is crucial to test for asymmetric employer learning.

For each respondent we construct the labour market history from the Work-History file which contains week-by-week longitudinal work records. For multiple job holders, we define the main job as the job for which the worker worked most during the week. We only consider the main job for each respondent and ignore multiple jobs. We match to this data set the educational histories of each respondent, i.e. his enrollment status as well as his highest grade completed.

A major problem in the NLSY79 concerns the type of employment relations one should consider in the analysis. We consider only jobs after the respondent has made a transition from school to work. Such a transition is difficult to define since working while enrolled at school is very common in the US. See appendix E for our definition.

Our sample is created as follows. There are 6398 men in the NLSY79. We do not consider respondents belonging to the military sample. This leaves us with 5574 individuals. There are 236 respondents who never made a transition from school to

¹²It is, however, easy to show that in a model without unemployment and permanent non-pecuniary job characteristics, movers earn a higher wage, but are as able as stayers under symmetric learning.

work according to our definition. This leaves us with 5238 individuals. We drop individuals who entered the labour market before 1978 as detailed information about weeks worked and employers worked for is only available from January 1978. We lose 1547 individuals because of this. We also delete all part-time jobs (< 30 hours a week), jobs without pay and jobs of the self-employed. Finally, we delete unreasonable wage information -see appendix E for the definition. We lose 32 individuals and 10034 wage observations because of this. We end up with 3659 individuals and 41260 wage observations.

As Farber and Gibbons (1996) and Altonji and Pierret (2001), we use the Armed Forces Qualification Test (AFQT) as a variable that is correlated with ability but not observed by employers. The AFQT provides a summary measure for basic literacy and numeracy skills. It is generally seen as a good indicator for workers' cognitive ability, but less so for workers' communication skills, etc. Our tests are therefore best understood as tests for employer learning about workers' cognitive ability. In our sample the AFQT- score is missing for 154 individuals, and 1475 wage observations. Our final sample thus consists of 3505 individuals and 39785 wage observations.

We define a "mover" as a worker who has *permanently* left his job. Workers who have been laid off and return to their previous firm are considered as stayers. We do this because the employing firm has essentially the same information about these workers' productivity as other incumbent firms, and potentially superior information to other outside firms. We consider a transition as a job-to-unemployment transition if the respondent reported that he was actively searching for a job between two employment spells. If a respondent did not report having actively looked for a job, we consider the transition as a job-to-job transition independently of the length of the non-employment spell. We have repeated the empirical analysis using an alternative definition for a job-to-unemployment transition. According to this definition, a transition is considered as a job-to-unemployment transition if the worker started the new job 4 weeks or later after his old job ended independently of whether the worker was actively looking for a job. This does not qualitatively change our results.

We perform all tests separately by education. We distinguish three education

groups: high school dropouts, high school graduates, and university graduates. High school graduates are defined as workers with at least a high school diploma (including GED), but without a bachelor’s degree. Results for high school graduates with and without some college education are very similar so that we decided not to distinguish between these two groups.

Table 2.3 reports the means and standard deviations of the most important variables used for the analysis.

Table 2.3: Means and standard deviations of selected variables

	high school dropouts N=5647	high school graduates N=25448	university graduates N=8690
log-wage	1.4260 (0.4013)	1.5879 (0.4387)	1.9806 (0.5127)
experience	5.1057 (3.6629)	5.5279 (3.8682)	5.2185 (3.5766)
tenure	1.8824 (2.3201)	2.3652 (2.6879)	2.725 (2.6832)
AFQT	13.5367 (13.3827)	34.4782 (24.9399)	73.7986 (22.9785)
prop. job-to-job transitions	0.20467 (0.4035)	0.1949 (0.3962)	0.1875 (0.3903)
prop. job-to-unemp. transitions	0.2277 (0.4194)	0.1686 (0.3746)	0.0895 (0.2855)

2.5 Empirical evidence

Our testing strategy proceeds in two steps. We first compare the AFQT- score of stayers, job-to-job and job-to-unemployment movers (section 2.5.1). Our second test is a test for whether easy- and hard-to-observe variables have the same or different impact on wages for movers and stayers (section 2.5.2). We finally discuss whether the results are also compatible with alternative explanations, such as a human capital model (section 2.5.3).

2.5.1 Stayers, job-to-job, and job-to-unemployment movers

In this section, we compare the average ability of three groups of workers: stayers, job-to-job and job-to-unemployment movers. Under symmetric learning, it is only job-to-unemployment movers who are less able than stayers (proposition 1). Under

asymmetric learning, on the other hand, it is job-to-job movers, and possibly also job-to-unemployment movers, who are less able than stayers (proposition 2). We test these implications by comparing the AFQT- score of the three groups of workers¹³. This strategy has several advantages to comparing the wage of stayers, job-to-job and job-to-unemployment movers. First, a higher ability of stayers may not necessarily result in a higher wage for stayers if non-pecuniary characteristics are not transitory or not drawn from an extreme value distribution. Second, differences in wages do not necessarily reflect differences in ability. For instance, job-to-unemployment movers may earn a lower wage than stayers not because they are less able but because they lost valuable firm- and match-specific capital.

We first regress the AFQT- score only on two dummy variables that classify a worker as a job-to-job or as a job-to-unemployment mover, respectively, with stayers as the base group. The AFQT- score is measured as the respondents' quantile in the test score distribution (i.e. a score of 75 means that 25 % of the test takers have a higher AFQT- score). In the regression for all education groups, we additionally control for workers' education. The hypothesis of asymmetric learning implies that the coefficient on the job-to-job mover is negative. Our null hypothesis therefore is

$$\begin{aligned}
 H_0 & : \beta^{\text{job-to-job}} = 0 \text{ (no asymmetric learning),} \\
 H_1 & : \beta^{\text{job-to-job}} < 0 \text{ (asymmetric learning).}
 \end{aligned}$$

Table 2.4, panel A, reports results. Standard errors are based on White/Huber standard errors that are clustered at the individual level. First note that distinguishing between job-to-job and job-to-unemployment movers is important: Job-to-unemployment movers tend to have lower test scores than job-to-job movers for all education groups. Comparing the average ability of stayers with that of job-to-unemployment movers -as suggested by Gibbons and Katz test for asymmetric learning- may thus give misleading results.

¹³An alternative test is to estimate the effect of ability on the probability of staying, moving from job-to-job, and becoming unemployed. We have also estimated a multinomial logit model with the AFQT score as an explanatory variable. The conclusions drawn from these estimates are the same as those from table 2.4.

Among all education groups, job-to-job movers are less able than stayers by a modest, but significant amount. We therefore reject the null hypothesis of no asymmetric learning. The average AFQT- score of job-to-job movers is about 1 percentage points lower than stayers, compared to 4.5 percentage points for job-to-unemployment movers. However, breaking down the analysis by workers' education reveals important differences across education groups. Contrary to the hypothesis of asymmetric learning, we find a significantly higher AFQT- score for job-to-job movers than for stayers among high school dropouts. Among high school graduates, job-to-job movers are less able than stayers by a small and insignificant amount. Among university graduates, in contrast, job-to-job movers have considerably lower test scores than stayers and the null hypothesis of no asymmetric learning is clearly rejected.

Table 2.4: Average AFQT score of job-to-job and job-to-unemployment movers

	all N=39785	hs dropouts N=5647	hs graduates N=25448	uni graduates N=8690
A job-to-job	-0.9385** (0.4116)	1.6015** (0.6691)	-0.7128 (0.5680)	-3.2240** (0.7566)
job-to-un.	-4.5286** (0.4676)	-0.6811 (0.5829)	-5.6134** (0.6248)	-3.7142** (1.2161)
B job-to-job	-0.7867** (0.3681)	1.3695** (0.6630)	-0.5597 (0.6630)	-2.5114** (0.6549)
job-to-un.	-2.6856** (0.4180)	-0.4982 (0.5683)	-3.0919** (0.5509)	-2.2813** (1.0369)
C job-to-job	-0.3556 (0.3446)	1.3987** (0.6376)	-0.1265 (0.4604)	-1.9867** (0.6009)
job-to-un.	-1.4654** (0.3885)	-0.3009 (0.5516)	-1.6103** (0.5089)	-1.0497 (0.9142)
D occ. and ind. sorting (p-value)	0.000	0.000	0.000	0.000

Dependent variable: Unadjusted 1989 AFQT-score. Base category: Stayers. The first panel controls for education only (pooled regression). The second panel additionally controls for the age of the respondent at time of test (6 dummies), time effects and race. The third panel additionally includes 12 industry and 10 occupation dummies as well as a dummy when the occupation or industry classification is missing. White/Huber standard errors clustered at the individual level in parentheses. H_0 tests for whether the industry and occupation dummies are jointly equal to 0.

The correlation between the AFQT- score and mobility may simply be due to omitted variables that are correlated with ability and also affect mobility, for reasons

not captured by our model. For instance, respondents who were older when they took the test do on average have higher test scores. Older workers may also be less mobile. In panel B we additionally control for year effects, race, and the respondent's age at the time the test was taken. The additional controls weaken the correlation between the AFQT- score and job-to-job as well as job-to-unemployment mobility for all education groups. Our overall conclusions, however, are unchanged.

In a third step we also control for workers' industry and occupation affiliation (table 2.4, panel C). Suppose that industries and occupations differ with respect to career opportunities. Some industries and occupations offer more stable jobs, and more able workers sort into more stable industries¹⁴. Unconditional on industry and occupation, such a sorting model provides an alternative explanation for a negative correlation between ability and mobility. Controlling for workers' occupation and industry affiliation considerably weakens the correlation between ability and job-to-job as well as job-to-unemployment mobility for all education groups. This shows that sorting according to ability into industries and occupations is important. The hypothesis that the occupation and industry dummies are jointly equal to zero is strongly rejected for all education groups (table 2.4, panel D). Our overall conclusions, however, are unchanged: Job-to-job movers are of lower ability than stayers among university graduates only. Interestingly, for this group of workers it is only job-to-job movers who are significantly less able than stayers, but not job-to-unemployment movers. The hypothesis that the test score of job-to-job and job-to-unemployment movers is the same, however, cannot be rejected.

To sum up, we find no evidence for a lower ability of job-to-job movers than of stayers among high school dropouts and high school graduates. For university graduates, in contrast, job-to-job movers are significantly less able on average than stayers. The empirical evidence is therefore supportive of asymmetric employer learning for university

¹⁴See Neal (1998) for a model along these lines. In his model more able workers choose to work in industries and occupations that offer more specific skills. His model predicts a negative correlation between ability and mobility *unconditional* on workers' industry and occupation affiliation. This correlation should disappear after controlling for workers' occupation and industry affiliation.

graduates, but not for high school dropouts and graduates. Note that this is consistent with our assumption that output of better educated workers is more sensitive to ability.

2.5.2 Learning by experience or tenure?

In this section we test whether the workers' AFQT-score and education have the same or different impact on wages for movers and stayers. To facilitate a comparison with results by Altonji and Pierret (2001) and Farber and Gibbons (1996), we use the age-adjusted test score and normalize the test score to have zero mean and a standard deviation of 1¹⁵. Standard errors are based on White/Huber standard errors that are clustered at the individual level. In the first column of table 2.5 we reproduce Altonji

Table 2.5: Learning by experience or tenure?

	1	2	3
AFQT	0.0664** (0.0044)	0.0641** (0.0043)	0.0667** (0.0178)
exp.	0.1224** (0.0050)	0.0923** (0.0053)	0.0929** (0.0059)
exp. · AFQT	0.0074** (0.0006)	0.0066** (0.0008)	0.0069** (0.0032)
exp. · hs. gr.	-0.0055** (0.0026)	-0.0092** (0.0029)	-0.0093** (0.0039)
exp. · uni. gr.	-0.0070** (0.0032)	-0.0072** (0.0673)	-0.0040 (0.0048)
ten.		0.0719** (0.0050)	0.0722** (0.0068)
ten · AFQT		0.0007 (0.0011)	0.0010 (0.0055)
ten · hs. gr.		0.0086** (0.0030)	0.0082 (0.0056)
ten · uni. gr.		0.0059 (0.0050)	-0.0013 (0.0067)
exp. · AFQT · hs. gr.			0.0007 (0.0035)
exp. · AFQT · uni. gr.			-0.0033 (0.0037)
ten. · AFQT · hs. gr.			-0.0019 (0.0056)
ten. · AFQT · uni. gr.			0.0048 (0.0061)

N=39785. Column 1 repeats Altonji and Pierret's analysis. Column 2 adds tenure-AFQT as well as tenure-education interactions. Column 3 adds experience-AFQT-education and tenure-AFQT-education interactions to allow for differences across education groups. All regressions control for education, time dummies, interactions between education and time dummies, experience squared, and cubed, tenure squared and cubed and race. White/Huber standard errors clustered at the individual level in parentheses.

¹⁵Farber and Gibbons (1996) use the wage level, whereas Altonji and Pierret (2001) use the logarithm of the wage level as the dependent variable. Our results are qualitatively unchanged if we use the wage level instead of the log-wage as the dependent variable.

and Pierret’s tests for symmetric learning. We confirm their finding that the coefficient on the AFQT- score increases with experience. Furthermore, consistent with Altonji and Pierret’s results, returns to experience are lower for high school and university graduates than for high school dropouts when the AFQT- score is included in the wage regression¹⁶.

The second column in table 2.5 additionally includes interaction terms between tenure and the AFQT- score and education, respectively. Under pure symmetric learning only the coefficient on the experience-test score interaction should be positive, while under pure asymmetric learning only the coefficient on the tenure-test score interaction should be positive. We specify our null hypothesis as

$$\begin{aligned} H_0 & : \beta^{\text{AFQT}\cdot\text{tenure}} = 0 \text{ (no asymmetric learning),} \\ H_1 & : \beta^{\text{AFQT}\cdot\text{tenure}} > 0 \text{ (asymmetric learning).} \end{aligned}$$

The coefficient on the tenure-AFQT- score interaction is close to 0 and not significantly different from zero. Moreover, the inclusion of the tenure interactions have little effect on the coefficient on the experience-AFQT- score and experience-education interactions. We therefore cannot reject the null hypothesis of no asymmetric learning, and learning appears to be largely symmetric.

Symmetric learning further implies that the coefficient on the interaction term between experience and education is negative if the AFQT- score is included in the wage regression. Asymmetric learning, in contrast, implies a negative interaction term between tenure and education. Our null hypothesis is

$$\begin{aligned} H_0 & : \beta^{\text{hs.grad}\cdot\text{tenure}} = 0, \beta^{\text{uni.grad}\cdot\text{tenure}} = 0 \text{ (no asymmetric learning),} \\ H_1 & : \beta^{\text{hs.grad}\cdot\text{tenure}} < 0, \beta^{\text{uni.grad}\cdot\text{tenure}} < 0 \text{ (asymmetric learning).} \end{aligned}$$

Contrary to the hypothesis of asymmetric learning, returns to tenure are higher for high school and university graduates than for high school dropouts. This difference is significant at a 5 % level for high school graduates only. On the other hand, returns to tenure are lower for university than for high school graduates. This difference is not

¹⁶Note that Altonji and Pierret (2001) use years of schooling instead of education groups.

significant at conventional levels. To conclude, results from column 2 are supportive of symmetric learning, but not of asymmetric learning

The specification in column 2 assumes that asymmetric information matters as much for university graduates as for high school graduates and dropouts. In particular, ability is assumed to have the same value at all jobs, i.e. $s = 0$. Results from the previous section indicate that there are important differences across education groups. Column 3 of table 2.5 therefore allows asymmetric information to matter more for better educated workers. It additionally includes interactions between schooling, the AFQT-score and experience and tenure, respectively. This is the specification implied by the model if the productivity of educated workers is more sensitive to ability, i.e. if $s > 0$. While none of the additional interactions is individually significant, they are jointly significant at a 5 % level (p-value: 0.0001). Again, there are important differences across education groups. For high school dropouts, the estimate for the tenure-AFQT interaction is close to zero (0.0010). For high school graduates, it is even negative (0.0010-0.0019=-0.0009). For university graduates, in contrast, the estimate is positive (0.010+0.0048=0.0058), and considerably higher than the estimate for the AFQT-experience interaction (0.0072-0.0033=0.0039). In line with our previous results, this is consistent with asymmetric learning for university graduates, but not for high school dropouts and graduates.

Furthermore, the additional interactions reduce the coefficient on the education-tenure interactions. In contrast to asymmetric learning, the return to tenure continues to be higher for high school graduates than for high school dropouts. In line with asymmetric learning, the return to tenure is lower for university graduates than for high school graduates. This difference is significant at a 5 % level (p-value=0.013).

Table 2.6 reports results separately by education. Compared to table 2.5, table 2.6 allows all set of controls to vary with education. Results are -not surprisingly- very similar to table 2.5. The coefficient on the interaction between tenure and the test score is positive and significant only for university graduates, but not for high school graduates and dropouts. Note that for university graduates, the coefficient on the experience-AFQT- score interaction is smaller than that on the corresponding tenure

interaction, but it is positive and significant. Hence, the hypothesis of *pure* asymmetric learning is also rejected for university graduates.

Table 2.6: Learning by experience or tenure? Results by education

	hs dropouts N=5647	hs graduates N=25448	uni graduates N=8690
AFQT	0.0610** (0.0273)	0.0510** (0.0092)	0.1051** (0.0096)
exp. · AFQT	0.0068* (0.0040)	0.0076** (0.0019)	0.0038** (0.0021)
ten · AFQT	0.0030 (0.0098)	-0.0008 (0.0029)	0.0057** (0.0029)

Same specification as table 2.5. White/Huber standard errors clustered at the individual level in parentheses.

Finally, observe that results from table 2.5 and 2.6 are broadly consistent with the single crossing property. The coefficient on the interaction between tenure and the AFQT- score is higher for university graduates than for high school graduates and dropouts, although this relationship is not monotone. Moreover, the coefficient on the AFQT- score is higher for university graduates than for high school graduates and dropouts. The hypothesis that the coefficient on the AFQT- score and the AFQT-tenure interaction is the same for all education groups is clearly rejected.

We have also estimated all regressions including interactions between the AFQT-score and time dummies. This reduces the interaction term between ability and experience, but has little impact on the other coefficients. Our overall conclusions are unchanged.

To sum up, the results from this second test are in line with those from the previous test. For high school dropouts and graduates, there is no evidence for asymmetric learning. For university graduates, in contrast, the data is supportive of the hypothesis of asymmetric learning. An important question is whether there are alternative explanations for these results. We next turn to this question. Since we find evidence in favor of asymmetric learning only for university graduates, we focus on this group of workers.

2.5.3 Alternative explanations

Are there alternative explanations for our findings that for university graduates job-to-job movers are less able than stayers and that the effect of the AFQT- score increases with tenure? A possible candidate is a human capital model in which more able workers accumulate more firm-specific human capital. As asymmetric learning, such as model predicts a lower ability not only of job-to-unemployment movers, but also of job-to-job movers. It also implies a positive coefficient on the AFQT- tenure interaction term. That a complementarity between ability and firm-specific human capital accumulation may be of potential concern is illustrated by several studies have found that better educated workers receive more training (e.g. Lynch (1992), Mincer (1988)). Altonji and Spetzler (1991) find that aptitude and achievement measures also have a positive impact on the probability of receiving training.

One way to distinguish the firm-specific human capital from the asymmetric learning model is to compare the returns to tenure by education groups. If more able workers accumulate more firm-specific human capital, we may expect better educated workers to do so, too. The firm-specific human capital model should therefore predict higher returns to tenure for better educated workers. Asymmetric learning, in contrast, predicts lower returns to tenure for better educated workers, conditional on the AFQT-score. Unfortunately, the empirical results are somewhat inconclusive in this respect. On the one hand, returns to tenure for high school graduates are significantly higher than those for high school dropouts. On the other hand, returns to tenure are significantly lower for university than for high school graduates (table 2.5, column 3). Higher returns to tenure for better educated workers should not be interpreted as clear evidence against asymmetric employer learning: If better educated workers acquire more firm-specific human capital we may observe higher returns to tenure for better educated workers even if there is asymmetric learning. It should also be noted that there are human capital models that predict lower returns to experience and tenure for better educated workers if more able workers accumulate more firm-specific and general human capital (e.g. Lange (2003)). An alternative way to distinguish a firm-specific human capital from an asymmetric learning model is to use the training data in the

Table 2.7: Average AFQT score of job-to-job and job-to-unemployment movers: The effect of training measures

	all N=39571	hs dropouts N= 5589	hs graduates N= 25303	uni graduates N= 8679
A job-to-job	-0.6142 (0.4080)	1.4942** (0.6302)	-0.3274 (0.5650)	-2.7979** (0.7474)
job-to-un.	-3.9814** (0.4590)	-0.5384 (0.5697)	-4.7867** (0.6114)	-3.2859** (1.2024)
ST	2.1647** (0.4309)	-1.4028 (1.4011)	3.2145** (0.5434)	0.9520 (0.8473)
$\sum ST$	1.9113** (0.4150)	3.0087** (1.5079)	2.8554** (0.5519)	-0.3000 (0.6479)
CT	2.2008** (0.5653)	6.5331** (2.5191)	3.2272** (0.7962)	-0.06023 (0.8357)
$\sum CT$	1.2866** (0.2953)	3.1549** (2.3282)	1.9905** (0.4720)	0.8694** (0.3942)
B job-to-job	-0.3990 (0.3644)	1.2761** (0.6196)	-0.1801 (0.4967)	-2.2813** (0.6345)
job-to-un.	-2.2300** (0.4123)	-0.3526 (0.5579)	-2.5333** (0.5423)	-1.7712** (0.6345)
ST	1.7179** (0.3921)	-1.8958** (1.3489)	2.3982** (0.4970)	1.3819* (0.7287)
$\sum ST$	1.3809 (0.3675)	3.1142** (1.4812)	2.0152** (0.4869)	-0.3939 (0.5879)
CT	2.4439** (0.4939)	5.2332** (2.3662)	3.4565** (0.6695)	0.6748 (0.7470)
$\sum CT$	1.5368** (0.2734)	3.0594** (2.3134)	1.8824** (0.4039)	1.5081** (0.3998)
C job-to-job	-0.1147 (0.3444)	1.3059** (0.5974)	0.0734 (0.4610)	-1.3703** (0.5946)
job-to-un.	-1.2152** (0.3868)	-0.2037 (0.5416)	-1.3387** (0.5058)	-0.5347 (0.9039)
ST	1.3663** (0.3913)	-1.6803 (1.3827)	1.9380** (0.4864)	0.9119 (0.7194)
$\sum ST$	1.1169** (0.3510)	2.7164* (1.4761)	1.6017** (0.4677)	-0.3830 (0.6008)
CT	1.6008** (0.4703)	4.9006** (2.4498)	2.5000** (0.6305)	-0.1007 (0.7073)
$\sum CT$	1.1291** (0.2585)	2.9723 (2.3309)	1.2304** (0.3757)	1.4023** (0.3718)

This table repeats the analysis of table , but additionally controls for whether the respondent received schol- and company based training (ST and CT, respectively) as well as the sum of school- and company-based training ($\sum ST$ and $\sum CT$, respectively). Dependent variable: Unadjusted 1989 AFQT-score. Base category: Stayers. The first panel controls for education only (pooled regression). The second panel additionally controls for the age of the respondent at time of test (6 dummies), time effects and race. The third panel additionally includes 12 industry and 10 occupation dummies as well as a dummy when the occupation or industry classification is missing. White/Huber standard errors clustered at the individual level in parentheses.

NLSY and attempt to account for different investments in firm-specific human capital. This strategy has several problems. First, in the NLSY respondents are consistently asked about formal training programs only. Human capital accumulation, however,

may be largely informal and reflect learning-by doing¹⁷.

Second, it is difficult to distinguish general training from firm-specific training. We use training that takes place at the employer (company training) as a proxy for firm-specific training, and training that takes place at school (school training) as a proxy general training. The exact definition can be found in appendix C. While this classification is clearly imperfect, there is some evidence that company training is more firm-specific than school training. One piece of evidence comes from Loewenstein and Spetzler (1999). In 1993, questions were asked about usefulness of the training program in doing the same kind of work at a different employer than at the current one. Overall, training appears to be largely general, with 63 % saying that all or almost all of the skills learned are useful at different employers. But workers report company training to be significantly more firm-specific than school training. Furthermore, consistent with the idea that company training is more firm-specific, Loewenstein and Spetzler (1997) find that workers receiving company training are less likely to leave their employer, while workers receiving school training are about as mobile as workers with no training¹⁸.

A third problem is that in the NLSY training questions have not been asked consistently. The most important change occurred in 1987. Before 1987, only information on training programs lasting one month or more was collected. After 1987, information on all training programs is available. The proportion of workers receiving training therefore jumps up in 1987.

Table 2.7 repeats the analysis of table 2.4. It compares the average AFQT- score of job-to-job and job-to-unemployment movers, controlling for several training measures. ST and CT are indicators of whether the worker has received school or company training, respectively. $\sum ST$ denotes the sum of school training programs workers have taken since they entered the labour market, and $\sum CT$ denotes the sum of company training programs workers have participated in since they started working for their current employer. In line with results by Altonji and Spetzler (1991), we find strong positive relationship between training measures and ability. Interestingly, this correla-

¹⁷The NLSY began asking questions on informal training in 1993.

¹⁸Lynch (1992) finds the same for workers receiving on-the-job and off-the-job training, respectively.

tion is considerably stronger for high school dropouts and graduates than for university graduates. The inclusion of the training measures uniformly weakens the correlation between the AFQT- score and job-to-job and job-to-unemployment mobility, indicating that human capital accumulation matters. However, the differences are not significant. Our overall conclusions are unchanged: Job-to-job movers have significantly lower test scores than stayers only among university graduates. Table 2.8 repeats the analysis

Table 2.8: Learning by experience or tenure? The effect of training measures

	1	2	3
AFQT	0.0634** (0.0043)	0.0598** (0.0177)	0.1031** (0.0108)
exp.	0.0911** (0.0053)	0.0919** (0.0059)	0.0949** (0.0115)
exp. · AFQT	0.0055** (0.0008)	0.0061** (0.0032)	0.0036* (0.0021)
exp. · hs. gr.	-0.0093** (0.0029)	-0.0098** (0.0032)	-
exp. · uni. gr.	-0.0054** (0.0037)	-0.0036 (0.0039)	-
ten.	0.0723** (0.0050)	0.0741** (0.0068)	0.0961** (0.0114)
ten · AFQT	0.0002 (0.0011)	0.0025 (0.0055)	0.0052* (0.0029)
ten. · hs. gr.	0.0074** (0.0030)	0.0054 (0.0056)	-
ten. · uni. gr.	0.0024 (0.0040)	-0.0054 (0.0066)	-
exp. · AFQT · hs. gr.		0.0002 (0.0066)	-
exp. · AFQT · uni. gr.		-0.0021 (0.0033)	-
ten. · AFQT · hs. gr.		-0.0038 (0.0056)	-
ten. · AFQT · uni. gr.		0.0023 (0.0060)	-
ST	0.0002** (0.0101)	0.0004 (0.0101)	0.0250 (0.0212)
$\sum ST$	0.0293** (0.0026)	0.0299** (0.0026)	0.0169** (0.0052)
CT	0.0571** (0.0102)	0.0578** (0.0102)	0.0443** (0.0175)
$\sum CT$	0.03101** (0.0031)	0.0306** (0.0030)	0.0180** (0.0049)

N=39571 for all education groups and N=8679 for university graduates. Column 1 repeats column 2, table 3, and additionally controls for whether the respondent received school- and company based training (ST and CT, respectively) as well as the sum of school- and company-based training ($\sum ST$ and $\sum CT$, respectively). Column 2 repeats column 3, table 3, and additionally controls for the same training measures. Column 3 repeats the analysis for university graduates only. All regressions control for education, time dummies, interactions between education and time dummies, experience squared, and cubed, tenure squared and cubed and race. White/Huber standard errors clustered at the individual level in parentheses.

of table 2.5 and 2.6. It analyses how the impact of education and the AFQT- score varies with experience and tenure, conditional on training measures. As expected,

both company and school training have a positive impact on wages. In column 1 and 2, the coefficients on the interactions between tenure and education goes down when training measures are included, but the pattern is the same. The third column uses university graduates only. Training measures have little impact on AFQT-tenure and AFQT-experience interactions.

To sum up, table 2.7 and 2.8 suggest that human capital accumulation matters, but our results are robust to the inclusion of training measures. Since the training measures in the NLSY are clearly imperfect, it is hard to rule out for sure that our results for university graduates are driven by heterogeneity in firm-specific human capital accumulation. There is, however, one further piece of evidence against the firm-specific human capital model. If more able workers acquire more firm-specific human capital, workers with a high wage growth today should experience a high wage growth tomorrow. Residuals from a within-job wage regression should therefore be positively correlated. Schoenberg (2004) analyses the time series properties of these residuals, using the NLSY. In line with the results by Topel (1991) and Topel and Ward (1992), there is no evidence for a positive correlation.

2.6 Conclusion

When workers enter the labour market, employers are likely to have only imperfect information about workers' productivity. The question of this paper is: Do both incumbent and recruiting firms have the same information about workers' productivity (symmetric learning), or do incumbent firms have superior information about workers' productivity (asymmetric learning)? The main contribution of this chapter is to derive new tests for the hypothesis of asymmetric employer learning. We further contribute to the literature on job mobility by analyzing the impact of ability and learning about ability on the worker's decision to switch jobs.

Our tests require a variable that is correlated with workers' productivity, but not observed by employers. We follow the existing literature and use the Armed Forces Qualification Test (AFQT) for this. The AFQT score provides a summary measure

for basic literacy and numeracy skills and is thus an indicator for workers' cognitive ability. Our tests are therefore best understood as tests for employer learning about workers' cognitive ability as opposed to about workers' non-cognitive ability, such as communication skills, etc. We perform all tests separately by education. Our model implies that asymmetric information matters more for better educated workers if the productivity of better educated workers is more sensitive to ability. We find some empirical support for this: The AFQT- score has a stronger impact on wages, in particular on wages for stayers, for university graduates than for high school graduates and high school dropouts.

Empirically, a distinction between education groups turns out to be important. For high school dropouts and graduates we find no evidence for asymmetric employer learning. Here, workers with low AFQT- scores are more likely to become unemployed, but not (significantly) more likely to move from job-to-job. Furthermore, in a wage regression the coefficient on the AFQT- score increases with experience, but not with tenure. In contrast, for university graduates we find consistent evidence in favor of asymmetric learning. Here, low ability workers are significantly more likely to move from job-to-job than high ability workers. Furthermore, the AFQT- score has a stronger impact on wages for stayers than for movers.

An important question is: Are their explanations other than asymmetric learning for the findings of university graduates? A model that is consistent with many of the implications of asymmetric employer learning is a model in which more able workers accumulate more firm-specific human capital. Such a model also implies a lower ability of job-to-job movers than of stayers as well as a positive coefficient on the interaction between tenure and the test score. The two models differ with respect to returns to tenure by education groups. If more able workers accumulate more firm-specific human capital, we expect better educated workers to do so, too. Returns to tenure should therefore be higher for better educated workers. Asymmetric learning, in contrast, predicts the opposite, conditional on the AFQT- score. We find that university graduates have lower returns to tenure than high school graduates, but not than high school dropouts. We have also attempted to account for different investments in firm-specific

human capital, using the training data available in the NLSY. Our results are robust to the inclusion of training measures.

We conclude that the way the market acquires new information about workers' productivity differs across education groups. This has important implications. First, our results imply that labour markets work more efficiently for high school dropouts and high school graduates than for university graduates. Furthermore, according to our results, firms earn -ex post- higher rents on university graduates than on high school dropouts and graduates. Productivity differences between university and high school graduates thus exceed wage differences. Our results also imply that firms have a stronger incentive to sponsor general training for university graduates than for high school graduates and dropouts. This may be one explanation for Lynch's (1992) finding that in the US better educated workers are more likely to receive company-related training.

2.7 Appendix

A Symmetric learning

A.1 Uniqueness of equilibrium

To simplify the notation, we drop the superscript k and the subscript i . The first order conditions of the maximization problems (2.1) and (2.2) are

$$\begin{aligned} w &= y - \frac{(e^{w/\lambda} + e^{v/\lambda} + e^{b/\lambda})\lambda}{e^{v/\lambda} + e^{b/\lambda}}, & \text{and} \\ v &= y - \frac{(e^{w/\lambda} + e^{v/\lambda} + e^{b/\lambda})\lambda}{e^{w/\lambda} + e^{b/\lambda}}. \end{aligned} \quad (2.5)$$

Clearly, $\frac{(e^{w/\lambda} + e^{v/\lambda} + e^{b/\lambda})\lambda}{e^{v/\lambda} + e^{b/\lambda}}$ is increasing in w . Hence, at the optimum the second derivative will be negative. The second order conditions for a maximum are thus satisfied. We have

$$\begin{aligned} \frac{dw}{dv} &= \frac{e^{w/\lambda}e^{v/\lambda}}{(e^{w/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{v/\lambda} + e^{b/\lambda})} < 1, & \text{and} \\ \frac{dv}{dw} &= \frac{e^{w/\lambda}e^{v/\lambda}}{(e^{w/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{w/\lambda} + e^{b/\lambda})} < 1. \end{aligned}$$

An increase in the competitor's wage leads to a less than proportionate increase in the own wage. Thus, the two upward-sloping curves intersect only once, and the equilibrium is unique.

A.2 Proof of proposition 1

(i) We first show that job-to-job movers are as able as stayers. To simplify the notation, we drop the superscript k . The argument is the same for workers with and without education. Incumbent and outside firms offer the same wage, i.e. $w_i = v_i$. Hence, the probability of staying is the same as the probability of moving from job to job for both low and high ability workers, i.e. $Pr(\text{job-to-job}|i) = Pr(\text{stay}|i)$. It follows that the proportion of workers with low ability is the same among stayers and job-to-job (jtj) movers, i.e.

$$\begin{aligned} Pr(L|\text{stay}) &= \frac{pPr(\text{stay}|L)}{pPr(\text{stay}|L) + (1-p)Pr(\text{stay}|H)} = \\ Pr(L|\text{jtj}) &= \frac{pPr(\text{jtj}|L)}{pPr(\text{jtj}|L) + (1-p)Pr(\text{jtj}|H)}. \end{aligned}$$

Hence, stayers are as able on average as job-to-job movers. It follows immediately that stayers and job-to-job movers have the same test score on average: The proportion of workers with a low test score is the same among stayers and job-to-job movers, i.e. $Pr(LT|job-to-job) = Pr(LT|stay)$.

(ii) The argument is trivial if the utility from unemployment is the same for low and high ability workers. Here, we show that the argument continues to hold if $b_H - b_L \leq a_H - a_L$. Using $w_i = v_i$, the probability that a high ability worker becomes unemployed equals $\frac{e^{b_H/\lambda}}{2e^{w_H/\lambda} + e^{b_H/\lambda}}$, while the probability that a low ability worker becomes unemployed equals $\frac{e^{b_L/\lambda}}{2e^{w_L/\lambda} + e^{b_L/\lambda}}$. Low ability workers are thus more likely to become unemployed iff

$$\begin{aligned} \frac{e^{b_L/\lambda}}{2e^{w_L/\lambda} + e^{b_L/\lambda}} &> \frac{e^{b_H/\lambda}}{2e^{w_H/\lambda} + e^{b_H/\lambda}} \\ \iff e^{w_H/\lambda} e^{b_L/\lambda} &> e^{b_H/\lambda} e^{w_L/\lambda} \\ \iff w_H - w_L &> b_H - b_L. \end{aligned}$$

Totally differentiating (2.5) with respect to ability, using $w_i = v_i$, yields

$$\frac{dw}{da} = \frac{(e^{w/\lambda} + e^{b/\lambda})^2 + \frac{\partial b}{\partial a} e^{w/\lambda} e^{b/\lambda}}{(e^{w/\lambda} + e^{b/\lambda})^2 + e^{w/\lambda} e^{b/\lambda}}.$$

Hence, as the denominator is positive, $\frac{dw}{da} \geq \frac{\partial b}{\partial a}$ iff $\frac{\partial b}{\partial a} \leq 1$: The wage difference between high and low ability workers exceeds the difference in the utility from unemployment iff $b_H - b_L \leq a_H - a_L$. It then follows that job-to-unemployment movers have lower test scores on average as stayers. The proportion of workers with a low test score among stayers and job-to-unemployment movers can be computed as

$$\begin{aligned} Pr(LT|stay) &= \frac{pq_L^{LT} Pr(stay|L) + (1-p)(1-q_L^{LT}) Pr(stay|H)}{Pr(stay)}, \text{ and} \\ Pr(LT|jtu) &= \frac{pq_L^{LT} Pr(jtu|L) + (1-p)(1-q_L^{LT}) Pr(jtu|H)}{Pr(jtu)}, \end{aligned}$$

where q_L^{LT} denotes the probability that a worker with low ability scores low in the test score. Taking the difference yields

$$\begin{aligned} Pr(LT|stay) - Pr(LT|jtu) &= \\ \frac{p(1-p)(2q_L^{LT} - 1)(Pr(stay|L) Pr(jtu|H) - Pr(stay|H) Pr(jtu|L))}{Pr(stay) Pr(jtu)} &< 0, \end{aligned}$$

as $q_L^{LT} > 0.5$. Hence, stayers have higher test scores on average as job-to-unemployment movers. The weaker is the correlation between the test score and ability, the lower is the difference between the average test score of stayers and job-to-unemployment movers. ■

A.3 Wage determination in the first period

In the first period firms observe workers' education but not their ability. Wages therefore depend on education, but not on ability. Let W^k and V^k , $k = u, e$, denote the first period wage offers of the two firms competing for uneducated and educated workers, respectively. Workers' utility from non-pecuniary characteristics at the two firms and when unemployed are represented by M_1, M_2 , and M_0 , respectively. In the first period there is symmetric information with respect to firms and workers, i.e. neither firms nor workers observe workers' ability. Let b^k denote worker's utility from unemployment in the first period. (As workers do not know their type, it may vary with education but not with ability). Workers and firms learn about workers' ability only during work. When the worker chooses to remain unemployed in the first period, she gets

$$\begin{aligned} U_1^k(\text{unemployment}) &= b^k + M_0 + E[\max(b^k + m_0, w^k + m_1, v^k + m_2)] \\ &: = b^k + M_0 + E_u^k, \quad k = u, e. \end{aligned}$$

Workers' utility in the first period when working at the first firm is

$$\begin{aligned} U_1^k(\text{work}) &= W^k + M_1 + p^k E[\max(b_L^k + m_0, w_L^k + m_1, v_L^k + m_2)] \\ &\quad + (1 - p^k) E[\max(b_H^k + m_0, w_H^k + m_1, v_H^k + m_2)] \\ &: = W^k + M_1 + E_w^k, \quad k = u, e. \end{aligned}$$

Workers' utility when working at the second firm can be derived similarly. The probability that the first firm attracts the worker thus is

$$\Pr(\text{work for firm 1}) = \frac{e^{(W^k + E_w^k)/\lambda}}{e^{(b^k + E_u^k)/\lambda} + e^{(W^k + E_w^k)/\lambda} + e^{(V^k + E_v^k)/\lambda}}.$$

Due to non-pecuniary characteristics, firms make positive profits in the second period. They take this into account in the first period. Let Π^k denote firms' profit. In the first period the first firm maximises

$$\begin{aligned} \max_{W^k} & \frac{e^{(W^k + E_w^k)/\lambda}}{e^{(b^k + E_u^k)/\lambda} + e^{(W^k + E_w^k)/\lambda} + e^{(V^k + E_v^k)/\lambda}} (p^k y_L^k + (1 - p^k) y_H^k + \beta \Pi^k - W^k), \\ k = & u, e; \end{aligned}$$

while the second firm maximises

$$\max_{V^k} \frac{e^{(V^k + E_w^k)/\lambda}}{e^{(b^k + E_u^k)/\lambda} + e^{(W^k + E_w^k)/\lambda} + e^{(V^k + E_w^k)/\lambda}} (p^k y_L^k + (1 - p^k) y_H^k + \beta \Pi^k - V^k),$$

$k = u, e.$

In equilibrium we have $V^k = W^k$. The probability of remaining unemployed is the same for low and high ability workers. In the first period workers in employment are thus as able on average as unemployed workers, and unemployment is not informative about workers' ability.

B Asymmetric learning

B.1 Uniqueness of equilibrium

To simplify the notation, we will drop the superscript k . The argument is the same for workers with and without education. The first order conditions of the maximization problem (2.3) and (2.4) are

$$\begin{aligned} w_i &= y_i - \lambda \frac{(e^{w_i/\lambda} + e^{v/\lambda} + e^{b/\lambda})}{e^{v/\lambda} + e^{b/\lambda}}, \quad i = L, H; \quad \text{and} \quad (2.6) \\ v &= \frac{\frac{pe^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2} y_L + \frac{(1-p)e^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2} y_H}{\frac{pe^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2} + \frac{(1-p)e^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2}} \\ &\quad - \lambda \frac{\frac{pe^{v/\lambda}}{e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda}} + \frac{(1-p)pe^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda}}}{\frac{pe^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2} + \frac{(1-p)e^{v/\lambda}(e^{w_L/\lambda} + e^{b/\lambda})}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})^2}} \\ &= \frac{pJ_L(1 - J_L)y_L + (1 - p)J_H(1 - J_H)y_H}{pJ_L(1 - J_L) + (1 - p)J_H(1 - J_H)} \\ &\quad - \lambda \frac{pJ_L + (1 - p)J_H}{pJ_L(1 - J_L) + (1 - p)J_H(1 - J_H)}, \end{aligned}$$

where J_L and J_H denotes the probability that a low and a high ability worker move from job-to-job, respectively. As $\frac{e^{w_i/\lambda} e^{v/\lambda}}{(e^{w_i/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{v/\lambda} + e^{b/\lambda})}$ is increasing in w_i , the objective function of the incumbent firm is quasi-concave. The objective function of the outside firm is thus a linear combination of two quasi-concave functions. Hence, the second order conditions for a maximum will be satisfied. We have

$$\frac{dw_i}{dv} = \frac{e^{w_i/\lambda} e^{v/\lambda}}{(e^{w_i/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{v/\lambda} + e^{b/\lambda})} < 1.$$

An increase in v by one unit increases w_i by less than a unit. Furthermore,

$$\begin{aligned}
& \{p(1-p)(y_H - y_L)(J_L(1 - J_L)(1 - 2J_H)\frac{\partial J_H}{\partial w_H} dw_H \\
& -p(1-p)(y_H - y_L)(J_H(1 - J_H)(1 - 2J_L)\frac{\partial J_L}{\partial w_L} dw_L \\
& +p(1-p)\lambda J_H(1 - 2J_L)\frac{\partial J_L}{\partial w_L} dw_L - p(1-p)\lambda J_H(1 - J_H)\frac{\partial J_L}{\partial w_L} dw_L \\
& +p(1-p)\lambda J_L(1 - 2J_H)\frac{\partial J_H}{\partial w_H} dw_H - p(1-p)\lambda J_L(1 - J_L)\frac{\partial J_H}{\partial w_H} dw_H \\
& -p^2\lambda J_L^2\frac{\partial J_L}{\partial w_L} dw_L - (1-p)^2\lambda J_H^2\frac{\partial J_H}{\partial w_H} dw_H\} \\
dv = & \frac{\hspace{10em}}{\hspace{10em}} \\
& \{(pJ_L(1 - J_L) + (1-p)J_H(1 - J_H))^2 \\
& -p(1-p)(y_H - y_L)(J_L(1 - J_L)(1 - 2J_H)\frac{\partial J_H}{\partial v} \\
& +p(1-p)(y_H - y_L)(J_H(1 - J_H)(1 - 2J_L)\frac{\partial J_L}{\partial v} \\
& -p(1-p)\lambda J_H(1 - 2J_L)\frac{\partial J_L}{\partial v} + p(1-p)\lambda J_H(1 - J_H)\frac{\partial J_L}{\partial v} \\
& -p(1-p)\lambda J_L(1 - 2J_H)\frac{\partial J_H}{\partial v} + p(1-p)\lambda J_L(1 - J_L)\frac{\partial J_H}{\partial v} \\
& +p^2\lambda J_L^2\frac{\partial J_L}{\partial v} + (1-p)^2\lambda J_H^2\frac{\partial J_H}{\partial v}\}
\end{aligned} \tag{2.7}$$

If an increase in w_L and w_H by one unit increases v by less than one unit, then the equilibrium will be unique. Note that $\frac{\partial J_i}{\partial v} = \frac{1}{\lambda} J_i(1 - J_i)$. Normalize $y_H - y_L$ to λ . It can then be verified that the denominator in (2.7) is positive. Also observe that $\frac{\partial J_i}{\partial v} \geq |\frac{\partial J_i}{\partial w_i}|$. It can then be checked that an increase in w_L and w_H by one unit leads to an increase of v by less than a unit. The equilibrium is thus unique.

B.2 Proof of proposition 2

(i) We drop the subscript k . The argument is the same for both education groups. Job-to-job movers are less able than stayers if $w_H > w_L$. If $s = 0$ and $a_L = a_H$, we have $w_L = w_H$. Hence, if $\frac{dw_H}{da_H} > \frac{dw_L}{da_H}$, then $w_L = w_H$. We have

$$\frac{dw_H}{da_H} = \frac{(e^{v/\lambda} + e^{b/\lambda})^2 + e^{w_H/\lambda} e^{v/\lambda} \frac{dv}{da_H}}{(e^{w_H/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{v/\lambda} + e^{b/\lambda})}, \tag{2.8}$$

while

$$\frac{dw_L}{da_H} = \frac{e^{w_L/\lambda} e^{v/\lambda} \frac{dv}{da_H}}{(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})(e^{v/\lambda} + e^{b/\lambda})}.$$

It can be shown that $\frac{dw_H}{da_H} > \frac{dw_L}{da_H}$.

We next show that the adverse selection is stronger for better educated workers, holding the proportion of low ability workers fixed. The difference between the productivity of stayers and job-to-job movers equals

$$E[y^k|\text{stay}] - E[y^k|\text{jtj}] = (y_H^k - y_L^k)(\Pr(L|\text{jtj},k) - \Pr(L|\text{stay},k)),$$

where $\Pr(L|\text{jtj},k) = \frac{p(e^{w_H/\lambda} + e^{v/\lambda} + e^{b/\lambda})}{p(e^{w_H/\lambda} + e^{v/\lambda} + e^{b/\lambda}) + (1-p)(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})}$ and $\Pr(L|\text{stay},k) = \frac{pe^{w_L/\lambda}(e^{w_H/\lambda} + e^{v/\lambda} + e^{b/\lambda})}{pe^{w_L/\lambda}(e^{w_H/\lambda} + e^{v/\lambda} + e^{b/\lambda}) + (1-p)e^{w_H/\lambda}(e^{w_L/\lambda} + e^{v/\lambda} + e^{b/\lambda})}$ is the proportion of low ability workers among job-to-job movers and stayers, respectively, for educated and uneducated workers. An increase in s clearly leads to an increase in $y_H^k - y_L^k$. Furthermore, it can be shown that an increase in s leads to an increase in $\Pr(L|\text{jtj},k)$ and a decrease in $\Pr(L|\text{stay},k)$.

(ii) This follows from $w_H > w_L$. ■

B.3 Proof of proposition 3

(i) Suppose the test score is perfectly correlated with ability, i.e. $q_L^{LT} = 1$. Then the coefficient on the test score identifies for workers without education who stay with their employer $w_H^u - w_L^u > 0$. If the test score is imperfectly correlated with ability, the coefficient on the test score identifies for workers without education

$$E[w^u|HT] - E[w^u|LT] = (w_H^u - w_L^u)(\Pr(L|LT, u, \text{stay}) - \Pr(L|HT, u, \text{stay})) = \frac{(w_H^u - w_L^u)p^u(1-p^u)\Pr(\text{stay}|L, u)\Pr(\text{stay}|H, u)(2q_L^{LT} - 1)}{\Pr(\text{stay}, LT, u)\Pr(\text{stay}, HT, u)} > 0.$$

Here, the condition HT and LT denote that the worker scored high or low, respectively, in the test score. Due to approximation error, the coefficient on the test score is attenuated towards zero.

(ii) Suppose the test score is perfectly correlated with ability. If the test score is included in the wage regression, the coefficient on worker's education identifies for stayers with low ability $w_L^e - w_L^u$. For movers returns to education equal $v^e - v^u$, independently of whether the test score is included in the wage regression. Since the proportion of low ability workers as well as the productivity of high ability workers affects outside wages, it also affects wage offers of low ability workers. If $p^u = p^e$, and $s = 0$, then $v^e = v^u$ and $w_L^e = w_L^u$. Hence, if for low ability workers s increases and p^e decreases outside wage offers by more than inside wage offers, returns to education will be higher for movers than for stayers, i.e. $v^e - v^u > w_L^e - w_L^u$.

Totally differentiating (2.6) with respect to s and p^e yields

$$\frac{dw_L^e}{ds} = \frac{e^{w_L^e/\lambda} e^{v^e/\lambda} \frac{dv^e}{ds}}{(e^{w_L^e/\lambda} + e^{v^e/\lambda} + e^{b/\lambda})(e^{v^e/\lambda} + e^{b/\lambda})} \leq \frac{dv^e}{ds}, \quad \text{and}$$

$$\left| \frac{dw_L^e}{dp^e} \right| = \left| \frac{e^{w_L^e/\lambda} e^{v^e/\lambda} \frac{dv^e}{dp^e}}{(e^{w_L^e/\lambda} + e^{v^e/\lambda} + e^{b/\lambda})(e^{v^e/\lambda} + e^{b/\lambda})} \right| \leq \left| \frac{dv^e}{dp^e} \right|.$$

Hence, the coefficient on education is higher for movers than for stayers, and the coefficient on the interaction between tenure and education is negative. ■

B.4 Proof of proposition 4

(i) Suppose the test score is perfectly correlated with ability. Under symmetric learning, in the second period the coefficient on the test score identifies for workers without education $w_H^u - w_L^u$, and for workers with education $w_H^e - w_L^e$. As $\frac{dw_H^e}{ds} > 0$, $w_H^e - w_L^e > w_H^u - w_L^u$.

(ii) Under asymmetric learning, the coefficient on the test score identifies for educated workers who stay with their employer $w_H^e - w_L^e$. For uneducated stayers, the coefficient identifies $w_H^u - w_L^u$. For movers, the coefficient on the test score is 0. We have to show that $w_H^e - w_L^e > w_H^u - w_L^u$. A productivity increase of high ability workers increases the outside wage offer, and consequently also the wage offer for low ability workers. From (2.8), we know that $\frac{dw_H^e}{ds} > \frac{dw_L^e}{ds}$. Consequently, a productivity increase of high ability workers increases the wage for high ability workers by more than the wage for low ability workers, and $w_H^e - w_L^e > w_H^u - w_L^u$. ■

B.5 Wage determination in the first period

In the first period information about workers' ability is symmetric. Hence, wage formation in the first period closely resembles wage formation under symmetric learning, and will not be further discussed here.

C Data

Definition of a school-to-work transition Our definition for a transition from school to work is a combination of Lynch's (1992) and Farber and Gibbons's (1996) definition. Lynch (1992) defines the transition from school to work as the year in which the respondent was

enrolled in school and that was followed by at least two years not enrolled in school. Farber and Gibbons (1996) look at whether the respondent is classified as working or not. A worker is classified as working when he has worked at least 26 weeks, and during these weeks at least 30 hours, since the last interview. We follow this classification, with the difference that we use calendar years instead of interview years. We do so because the time between two interviews varies from 9 to 40 months. According to Farber and Gibbons (1996), a transition from school to work, or rather from non-work to work, takes place when the worker was classified as non-working for at least one year, followed by at least two consecutive years classified as working. Our definition takes the minimum of Lynch's and Farber and Gibbon's definitions.

Wage collection At each interview respondents are asked whether they are still working for the employer they worked for at the last interview. If yes, they are asked whether their usual wage has changed. If no, they are asked whether the wage at the old job has changed before they switched jobs. They are also asked about the usual wage at the new job. We assume that wages refer to tenure and experience at the end of the employment spell. Results change very little if we assume that wages refer to tenure and experience at the beginning of each employment spell. We use the hourly wage rate which was computed by the Bureau of labour Statistics. Wages are deflated by the Consumer Price Index with 1995 as the base year.

Definition of unreasonable wage information We drop unreasonable wage information. We consider a wage observation as unreasonable if the hourly wage rate is below 1 \$ (this is well below the official minimum wage) and above 300 \$. We also drop a wage spell if the hourly wage increases (decreases) by at least 300 % in one period, followed by a 300 % decrease (increase) in the next period.

Definition of experience and tenure Actual experience is measured as weeks (divided by 52) spent in full-time employment after the transition from school to work. Part time employment, time spent as self-employed or working without pay, time spent unemployed as well as time spent out of the labour force is not counted. Actual tenure is measured as weeks (divided by 52) spent in full-time employment with the same employer.

Definition of education We distinguish between high school dropouts, high school graduates, and university graduates.

1. *High school dropouts* are workers without a high school diploma or GED.
2. *High school graduates* are workers with a high school diploma (including GED), but no bachelor's degree. High school graduates include college dropouts.
3. *College graduates* are workers with at least a bachelor's degree.

Definition of company and school training Each year, respondents in the NLSY were asked whether since the date of the last interview they attended any training program or on-the-job training. From 1987, information on up to 4 training programs was collected, independently of the duration of the training program. Before 1987, only information on up to 3 training programs was collected, for programs that lasted at least 1 month. Respondents were asked which categories best describes where they received the training. Before 1987, the categories are business college (1), nurses program (2), apprenticeship (3), vocational or technical institute (4), barber-beauty school (5), flight school (6), correspondence course (7), company training (8), or other (9). We classify category (8) as company training (CT), and the other categories as school training (ST). After 1987, the categories were as follows. Business school (1), apprenticeship (3), vocational or technical institute (4), correspondence course (7), formal company training run by employer or military (8), seminar training program at work run by someone other than employer (9), seminars or training programs outside of work (10), vocational rehabilitation center (11), other (12). We classify categories (8) and (9) as company training and the other categories as school training.

Chapter 3

Training and Unions¹

Standard human capital theory predicts that in perfectly competitive labour markets firms do not invest into general training of their workers, as the worker captures the full return to that investment (Becker 1964). Standard human capital theory also predicts that a minimum wage reduces on-the-job training, as it prevents workers from taking a wage cut during the training period (Hashimoto 1982). Unions should also reduce training investments if they compress wages, as workers do not capture the full return to training. Yet, there is evidence that suggests that firms provide and pay for general training². The evidence on the impact of minimum wage and unions on training is mixed. For the US, Neumark and Wascher (2001) report that workers subject to a minimum wage receive less training, whereas Acemoglu and Pischke (2001) find no such effect. For the UK, several studies indicate that workers covered by union agreements receive more training (e.g. Booth et al. (2003), Green et al. (1996)).

This chapter revisits the question how unions affect on-the-job training in the economy. We do this within the German apprenticeship system (GAS)³. The GAS is inter-

¹This chapter is joint work with Christian Dustmann.

²For instance, temporary help firms in the United States often provide and finance training programmes to new employees which is largely general in nature, such as computer and typing skills (Autor 2001). Another prominent example is the German apprenticeship system (e.g. Acemoglu and Pischke (1998)).

³See for instance Soskice (1994), Franz and Soskice (1995) and Winkelmann (1997) for a detailed description of the German apprenticeship system.

esting for two reasons. First, although apprentices are mostly trained in general skills, the system is, at least partly, financed by firms⁴. Second, collective bargaining plays an important role in Germany: Between 1996-1999 76 % of the West-German work force were covered by union agreements.

Our point of departure is the work by Acemoglu and Pischke (1999a, 1999b). Acemoglu and Pischke argue that firms have an incentive to sponsor general training if, due to labour market imperfections, the wage structure is compressed. Wage compression implies that training increases workers' productivity by more than workers' outside option, so that firms can increase profits by training. Then, contrary to the standard human capital theory, a minimum wage and union agreements may lead to more on-the-job training if they compress the wage structure. Based on this idea, we develop a simple model of unions and training, and test its empirical implications using matched firm-worker data.

Our model is intended to capture the specific features of the German collective bargaining system. In Germany, only firms that belong to an employer federation are legally obliged to pay at least the union wage. Membership in an employer federation is voluntary. Furthermore, union wages act as minimum wages and apply to all workers in a unionised firm. The most important difference between our and existing models of training and unions, such as Acemoglu and Pischke (1999b) and Booth et al. (1999), therefore is that in our model unionised and non-unionised firms co-exist. Unionised firms have to pay at least the union wage, but are allowed to pay a higher wage⁵. Non-unionised firms, in contrast, may pay whatever wage they want. Within this framework, we derive wage and training determination in unionised and non-unionised firms. Workers choose to work in the sector in which their utility is higher, while firms are indifferent between joining the unionised or non-unionised sector.

We analyse the firm's decision to train under the assumption that firms can only commit to training *provision*, but not to the *amount* of training. Consequently, workers are not willing to accept a wage cut to pay the training costs, and in a perfectly com-

⁴See studies by von Bardeleben et al. (1995) and Acemoglu and Pischke (1999b) for evidence.

⁵Our model therefore differs from union models such as Acemoglu and Pischke (1999b) and Booth et. al. (1999) who assume that firms are forced to pay the union wage.

petitive labour market, no worker will be trained. We show that in non-unionised firms wages are not compressed, so that non-unionised firms offer no training. In unionised firms, in contrast, the union wage compresses wages for workers with a productivity around the union wage. This induces unionised firms to train these workers. The role unions essentially play in our model is that they serve as a commitment device⁶: Unlike non-unionised firms, unionised firms commit to paying at least the union wage in the future, leading to wage compression in these firms.

There are explanations other than unions for firm-financed training. In a prominent paper, Acemoglu and Pischke (1998) suggest asymmetric employer learning as a reason for apprenticeship training in Germany. The key implication of their model is that trainees who leave the training firm (movers) are less able than trainees who stay with the training firm (stayers). They find this prediction confirmed in the data. Our model also has implications for the average productivity and wage of movers and stayers. We compare them with those of Acemoglu and Pischke's asymmetric learning model.

We derive two sets of empirical implications from our model. The first, and key, implication is that unionised firms offer more firm-financed training than non-unionised firms. We test this implication using firm level data that supplemented with information of the firm's workforce (obtained by matching all employees, drawn from administrative records, to each firms in the panel) and covers the period between 1996 and 1999. Our estimation strategy takes into account a nonrandom selection of workers and possibly firms into the unionised sector. Exploiting the changes in union status over time, we use a difference in difference estimator, combined with a linear matching estimator. We consistently find that unions increase training.

In a second step, we compare the empirical implications of Acemoglu and Pischke's asymmetric information model with those of our model. The second set of implications thus concerns worker selection into and from training firms. We show that the key implication of asymmetric information -i.e. a lower ability of movers than of stayers- is also compatible with our model. This is because unionised firms lay off all workers with

⁶A similar point has been made by Booth ()

a productivity below the union wage after training. Information on the union status of the training firm helps to distinguish between the alternative reasons for an adverse selection. Union agreements do not apply in non-unionised firms. Consequently, according to our union model we should observe an adverse selection only for workers trained in unionised firms, while asymmetric information predicts an adverse selection also for workers trained in non-unionised firms.

We test for an adverse selection using data from social security records, covering the period between 1975 and 1995. It allows constructing complete work and wage histories for apprentices before and after the training period. In line with the empirical evidence provided by Acemoglu and Pischke (1998), we find strong evidence for worker selection after the training period. We then use matched employee-employer data to compare the mover-stayer wage differential for workers trained in unionised and non-unionised firms. We find a higher stayer-mover wage differential not only for workers trained in unionised firms, but also for workers trained in non-unionised firms. However, the differential is twice as high for workers trained in unionised firms. This suggests that wage floors caused by unions contribute to an adverse selection of movers, but are not the only reason. We conclude that union agreements are an important reason for firm-financed training in Germany. However, we cannot rule out asymmetric information as an additional reason.

The structure of the paper is as follows. Section 3.1 describes the German collective bargaining system. Section 3.2 develops a model of employer-financed training. We begin with a base model and then incorporate union agreements into the model. Section 3.3 outlines the empirical implications and the empirical strategy. We then describe the data and results (section 3.4). We conclude by pointing out the policy and welfare implications of our findings (section 3.5).

3.1 Collective bargaining in Germany and outline of model

In Germany, negotiations between unions and employer federations take place at a regional and industry level, typically on an annual basis. The crucial feature of the German collective bargaining system is that not all firms are bound by union agreements. Only firms that belong to an employer federation (*Arbeitgeberverband*) are legally obliged to pay (at least) union wages. Membership in an employer federation is voluntary. In firms that belong to an employer federation, union agreements de facto apply to all employees, not only to union members. Workers who are union members but work in firms that do not belong to an employer federation are not entitled to union wages. The worker's union status is therefore irrelevant.

Firms that choose to not belong to an employer federation have two options. They can either engage in bilateral negotiations with the union or agree individual contracts with their employees. Between 1996 and 1999 49.95 % of firms in West Germany were members of an employer federation and bound by industry- and region-wide agreements. These firms employed 67.01 % of the West German work force. The proportion of firms which bilaterally negotiate with unions is comparatively small (6.39 %). 43.6 % of firms are neither bound by industry-wide or firm level agreements. Firms can change their union status. Between 1996 and 1999, 6.6 % of firms previously covered by union agreements left the employer federation, while 4.4 % of firms previously not covered by union agreements joined an employer federation⁷.

The most important outcome of the negotiation between unions and employer associations is the union wage. Union wages depend on easily observable worker characteristics, such as workers' skill, occupation, and experience. A different union wage applies for workers in apprenticeship training. Legally, union wages act as minimum wages, and firms may pay wages above the union wage. Payment above the union wage appears to be common. Bellmann et al. (1998) report that in 1997 49% of West German firms paid wages above tariff, although little is known about how many workers

⁷Own calculations based on the IAB-Betriebspanel.

in a firm receive wages above tariff. Besides wages, union agreements typically specify the weekly working time and overtime payments. Unions and employer federations do not bargain over employment, and unions do not have the power to prevent layoffs. Moreover, although firing costs in Germany are generally quite high, firms face no firing costs when laying off workers at the end of the apprenticeship.

Our union model is intended to capture these features of the German collective bargaining system. The crucial feature of our model is the coexistence of a unionised and non-unionised firms. The difference between the two types of firms is that unionised firms have to pay at least the union wage, and are allowed to pay a higher wage. Non-unionised firms, in contrast, may pay whatever wage they want. Workers differ with respect to their productivity, conditional on the region, industry and occupation they work in, but the same union wage applies to all of them. Firms can lay off workers at the end of the apprenticeship without incurring firing costs. Within this framework, we analyse wage and training determination of both unionised and non-unionised firms as well as the sorting of workers into unionised and non-unionised firms.

Note that although we model the union wage as a minimum wage, our model is different from a minimum wage model in several important aspects. Most importantly, in our model the union (minimum) wage is binding only in unionised firms, and firms and workers always have the option not to be unionised. Moreover, the union wage does not apply to workers in training.

3.2 A model of firm-financed training

This section sets up a simple model of firm-financed training. We first analyse firms' incentives to train when there are no union agreements (section 3.2.1). We then incorporate wage rigidities due to unions into the model (section 3.2.2). In order to focus on the impact of union agreements on training, we abstract from other reasons for wage compression and firm-financed training, such as complementarity between general and firm-specific skills, asymmetric information with respect to incumbent and outside firms (Acemoglu and Pischke (1998)), and asymmetric information with respect to workers

and firms (see Autor 2001 and Bhaskar and Holden 2002 models of this type).

3.2.1 Base model

There are many workers and firms. Firms and workers are risk-neutral. Firms maximise expected profits, workers maximise expected utility. There are two periods. The first period is the training period. There is no discounting.

Productivity and training costs Workers' productivity in period 2 depends on their (true) ability η as well as on the amount of training received in period 1, τ :

$$y = y(\tau, \eta).$$

We assume that $y(\tau, \eta)$ is strictly increasing, differentiable and concave in τ , with $\frac{\partial^2 y}{\partial \tau^2} < 0$. The return to training is higher for more able workers, i.e. $\frac{\partial^2 y}{\partial \tau \partial \eta} > 0$. The productivity of a worker in training is smaller than the productivity of an untrained worker by a constant k , which represents a fixed cost of training. There are also variable training costs which we denote by $c(\tau)$. The function $c(\tau)$ is strictly increasing, differentiable and convex, with $c(0) = c'(0) = 0$, $c''(\tau) > 0$. We further assume that the firm's production function exhibits constant returns to scale, i.e. the total productivity of a firm is equal to the sum of each worker's productivity.

Training decision In the first period, firms -as opposed to workers- decide how much training to offer to a worker. Training is continuous, and firms can condition their investment decision on workers' expected ability. We analyse the firm's decision to train under the assumption that firms can only commit to training *provision*, but not to the *amount* of training. We refer to this case as *limited commitment*. A justification for this assumption is that training is not easily verifiable by a third party, and can thus not be contracted upon. For our particular application - apprenticeship training schemes in Germany - the assumption that firms can commit to training provision, but not to the amount of training is reasonable. Trainees take centralised exams at the end of the apprenticeship training period and receive a certificate that is widely recognised. Hence, it is clearly verifiable whether a worker has received some training.

However, an important part of apprenticeship training takes place inside the firm, and is not easily observed by an outside party.

Information structure Workers' ability η is drawn from a normal distribution with mean $\bar{\eta}$ and variance σ_η^2 . In the first period firms and workers have the same information about workers' ability. Both parties receive a noisy signal $\tilde{\eta}$ about workers' ability. The signal equals

$$\tilde{\eta} = \eta + \varepsilon_\eta.$$

Firms and workers use the signal to update the belief about workers' ability. If ε_η is normally distributed with mean 0 and variance σ_η^2 , then the updated belief about the worker's productivity is also normally distributed (DeGroot (1970)). The updated belief about workers' productivity is a weighted average of the prior mean, $\bar{\eta}$, and the signal, $\tilde{\eta}$. We denote this updated belief by $\hat{\eta}$. Let $F_1(\eta|\hat{\eta})$ denote the ability distribution of a worker with expected ability $\hat{\eta}$ ⁸. In the second period both incumbent and outside firms symmetrically learn about workers' true ability. The assumption that firms perfectly learn about workers' ability is not essential for our results.

Mobility At the end of the training period workers decide whether to switch firms. As in Acemoglu and Pischke (1998), we assume that during the training period workers experience a utility shock θ . This shock captures the worker's ex post evaluation of her work environment. For example, it may reflect how well the worker gets along with her co-workers and supervisors. Only the worker, but not the firm, observes θ . We specify the worker's utility in period 2 at the incumbent firm, U^i , as a simple linear function of the incumbent firm's wage offer, w , and the utility from non-pecuniary job characteristics, θ :

$$U^i = w + \theta.$$

The worker's utility at outside firms is equal to the wage offer, v . The utility shock is drawn from a distribution with the cumulative distribution and probability density

⁸From DeGroot (1970), $E[\eta|\tilde{\eta}] = \frac{\bar{\eta}\sigma_\eta^2 + \tilde{\eta}\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\eta^2}$, and $V[\eta|\tilde{\eta}] = \frac{1}{\frac{1}{\sigma_\eta^2} + \frac{1}{\sigma_\eta^2}}$.

function G and g and support $[\underline{\theta}, \bar{\theta}]$. We assume that $G(\cdot)$ belongs to the family of log-concave distribution functions, i.e. $\frac{g(\theta)}{1-G(\theta)}$ is non-decreasing in θ . We also assume that the distribution of the utility shock neither depends on workers' ability nor on training.

Notice that the assumption that the upper support of G is positive, $\bar{\theta} > 0$, implies that *some* workers may stay with the employer even if offered a lower wage than outside firms. Consequently, firms make positive profits in the second period.

Wage determination Each period firms simultaneously make wage offers to the worker by maximising expected profits. Long-term wage contracts are not feasible.

Free entry condition Finally, we impose a free entry condition on firms: No firm earns positive profits in equilibrium at any point in time.

Timing The exact sequence of events is as follows.

1. At the beginning of period 1 firms and workers receive a noisy signal about workers' ability.
2. Firms offer a (first period) wage and choose whether and how much to train the worker.
3. Training/production takes place.
4. At the end of the training period all firms and workers get to know workers' ability. Incumbent and outside firms make a wage offer to the worker.
5. At the beginning of period 2 workers learn their utility shock. They decide whether to stay with or leave the training firm.
6. Production takes place.
7. At the end of the second period workers retire.

We solve the model backwards and begin with wage determination in the second period. We then turn to firms' incentives to train and wage determination in the first period.

Wage determination in the second period

Firms observe workers' ability so that the wage offer of incumbent and outside firms depend on ability. Wages also depend on training. Although outside firms may not be able to directly observe the worker's training level, in equilibrium they can infer the training level the incumbent firm has chosen in the first period, as they know the incumbent firm's maximisation problem⁹.

Let w denote the wage offer of the incumbent firm and v the worker's outside wage offer. Due to perfect competition in the outside market, outside firms bid up workers' wage until it equals the worker's (marginal) productivity, i.e. $v = y(\tau, \eta)$. Incumbent firms set wages by maximising expected profits, and trade off a higher chance of attracting the worker with a lower rent per worker. A worker stays with the training firm if the utility from staying, $w + \theta$, exceeds the utility from moving, $v = y$. Hence, the probability of staying is

$$Pr(\text{stay}) = Pr(\theta > y - w) = 1 - G(y - w).$$

If $\bar{\theta} > 0$, this probability is positive even if $w < y$: Some workers stay with the firm although they receive a lower wage offer from the incumbent than from outside firms. Incumbent firms maximise

$$\max_w (1 - G(y - w))(y - w).$$

From the first order condition, w satisfies

$$w = y - \frac{1 - G(y - w)}{g(y - w)}.$$

Log-concavity of G guarantees that the second order condition for a maximum is satisfied. Since workers stay with a positive probability with the incumbent firm even if they receive a higher outside wage offer, firms have monopsonic power and pay wages below productivity. It can be easily verified that $\frac{dw}{dy} = 1$: A productivity increase of one unit leads to a wage increase of the same magnitude. Hence, the wage offer of the

⁹Here we are implicitly assuming that outside firms observe the worker's expected ability in the first period (or her first period wage).

incumbent firm is equal to the worker's productivity minus a constant, Δ :

$$w := y - \Delta. \tag{3.1}$$

This is due to the assumption that the distribution of non-pecuniary job characteristics does not depend on workers' ability and training.

Training decision and wage determination in the first period

We now turn to the firm's training decision. Since commitment to training provision is limited, firms choose the training level by maximising future profits. They do not take into account the impact training has on the utility of the worker. From (3.1), firms earn a rent of Δ on each retained worker. A worker stays with the incumbent firm with probability $(1 - G(\Delta))$. Hence, the firm's profit in the second period equals $\Pi = (1 - G(\Delta))\Delta$. Clearly, it does not depend on training. Consequently, firms offer no training in equilibrium. Although non-pecuniary job characteristics lead to firms making positive profits, they do not compress the wage structure. We summarise

Proposition 1 *Under limited commitment firms offer no training.*

We finally have to analyse how wages are determined in the first period. Because of the free-entry condition firms make zero profits at any point of time. In the first period firms thus bid up the worker's wage until all rents are exhausted. As firms make positive profits in the second period, first period wages are higher than workers' expected productivity in the first period. Hence, there is front-loading. First period wages satisfy

$$W = E[y(0, \eta)|\hat{\eta}] + (1 - G(\Delta))\Delta. \tag{3.2}$$

3.2.2 Union agreements

We next incorporate wage rigidities due to unions into the training model. The model intends to capture the basic features of the German collective bargaining system (see section 3.1). Most importantly, a unionised and non-unionised sector coexist. Each

sector consists of many firms competing for workers. The difference between the two types of firms is that unionised firms have to pay at least the union wage, while non-unionised firms may pay whatever wage they want. We denote the union wage by \bar{w} . To simplify the notation, we assume that the same union wage applies for workers with and without an apprenticeship. Our results also hold if the union wage for workers with training exceeds that for workers without training, as long as training increases productivity more than the union wage. Union wages do not depend on ability, and do not apply for workers in training. Firms can layoff workers at the end of the apprenticeship without incurring any firing costs.

In Germany, firms can change their union status. In the theoretical analysis, we rule this out for simplicity. Our key result that unionised firms are more likely to train than non-unionised firms continues to hold if firms that were unionised in the past are more likely to be unionised in the future. This clearly holds in the data.

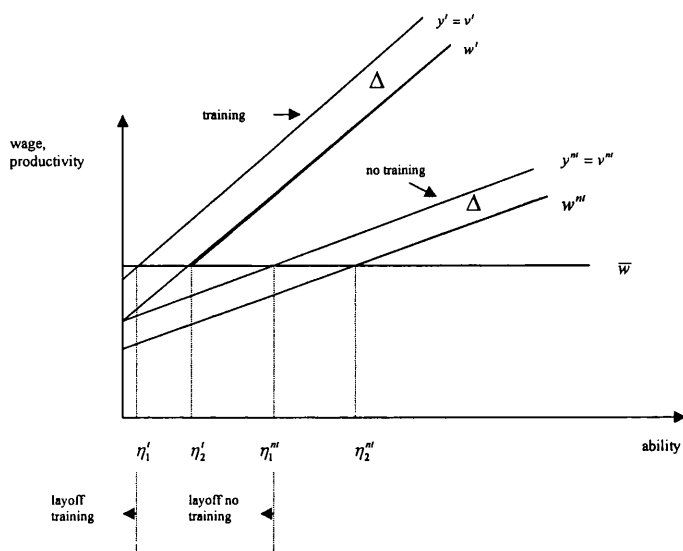
We commence by analysing wage determination in the second period. We then turn to firms' incentives to train and wage determination in the first period. Finally, we analyse the sorting of workers into the unionised and non-unionised sector, and give an interpretation why unions increase training.

Wage determination in the second period

Non-unionised firms First, consider non-unionised firms. Due to the free entry condition, unionised and non-unionised *outside* firms (i.e. firms that do not employ a particular trainee) continue to pay wages equal to workers' productivity. Consequently, union agreements do not affect wage determination of *incumbent* non-unionised firms: Since the worker's outside option is unaffected by union agreements and union agreements do not apply to non-unionised firms, non-unionised incumbent firms continue to offer a wage equal to the worker's productivity minus a constant Δ (see expression 3.1).

Unionised firms Union agreements affect wage offers of unionised incumbent firms. Figure 3.1 illustrates how wages are set in these firms. In the figure, we consider

Figure 3.1: Wage determination in unionised firms



untrained and trained workers. The vertical axis carries the wage and productivity of the worker, and the horizontal axis her revealed ability. Productivity and wages of untrained (trained) workers in the absence of any union agreement are indicated by the panels y^{nt} (y^t) and w^{nt} (w^t). From (3.1), they are equal to productivity minus a constant, Δ . The horizontal line indicate the union wage \bar{w} . It is useful to distinguish between three groups of workers.

1. Workers with productivity below the union wage. In the figure these are workers with ability below η_1^t if trained and η_1^{nt} if untrained. These workers are worse off due to union agreements. Unionised firms do not find it profitable to employ them. As there are no firing costs at the end of the apprenticeship, these workers are laid off. They find work in non-unionised firms and earn a wage equal to their productivity. Note that layoffs at the end of the apprenticeship occur because employers acquire new information about workers' ability during the training period. If unionised firms had known workers' ability in the first period, workers with an ability below η_1^{nt} (η_1^t) would not have been hired.
2. Next, consider workers with a productivity above the union wage, but whose

wage in the absence of union agreements falls below the union wage. In the figure, this refers to all workers with ability between η_1^t and η_2^t if trained, and η_1^{nt} and η_2^{nt} if untrained. These workers are better off due to unions, and earn a higher wage than they would in the absence of union agreements. Optimally, unionised incumbent firms would want to offer a wage below the union wage. As they are not allowed to do so, the best they can do is to offer just the union wage. Hence, workers with ability between η_1^t and η_2^t (η_1^{nt} and η_2^{nt}) are paid the union wage.

3. Finally, consider workers whose wage in the absence of union agreements exceeds the union wage. In the figure this applies to all workers with ability above η_2^t if trained, and η_2^{nt} if untrained. These workers are unaffected by union agreements. The union wage is not binding for these workers. They thus earn the same wage as in the absence of union agreements.

Note that the probability that a worker turns out to be less able than η_1^t (η_1^{nt}) depends on the worker's expected ability, $\hat{\eta}$.

Training decision and wage determination in the first period

We now turn to the training decision of unionised and non-unionised firms in the first period.

Non-unionised firms Non-unionised firms offer no training in equilibrium. This is because union agreements have no impact on wage determination of non-unionised firms. The future profit of non-unionised firms is thus the same as in the absence of union agreements, and is not increasing in training (proposition 1).

Unionised firms We next argue that training increases the future profits of unionised firms. This is best understood from figure 3.1. Consider a worker whose true ability is η_1^{nt} . Without training the firm would make zero profit on this worker. With training the worker's productivity is higher than the union wage \bar{w} , and the firm makes positive profits. More generally, training increases the rent on all workers with ability between

η_1^t and η_2^{nt} . Workers with ability below η_1^t are less productive than the union wage even after training. Workers with ability above η_2^{nt} are unaffected by union agreement even without training. Observe that this argument relies on firms making positive profits in the second period. Although non-pecuniary job characteristics are not sufficient to induce firms to sponsor training, they are necessary for unions to have an impact on training.

We now formalise this argument. We first derive an expression for the second period profit of unionised firms. Let $E[\Pi_u(\tau, \eta)|\hat{\eta}]$ denote the future (i.e. second period) profit on a worker with expected ability $\hat{\eta}$. Define η_1 as $y(\tau, \eta_1) = \bar{w}$, i.e. workers with ability below η_1 have a productivity below the union wage. Similarly, define η_2 as $y(\tau, \eta_2) = \bar{w} + \Delta$, i.e. workers with ability above η_2 are not affected by the union wage. Observe that η_1 and η_2 depend on the worker's training level. Unionised firms lay off workers with ability below η_1 and hence make zero profits on these workers. For workers with ability between η_1 and η_2 , unionised firms earn a rent of $y - \bar{w}$. These workers stay with the unionised firm after apprenticeship graduation with probability $1 - G(y - \bar{w})$. Finally, for workers with ability above η_2 , firms make a profit of $(1 - G(\Delta))\Delta$. Hence, unionised firms maximise

$$\begin{aligned} \max_{\tau} \quad & -c(\tau) + E[\Pi_u(\tau, \eta)|\hat{\eta}] = -c(\tau) + \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w}))(y - \bar{w})dF_1(\eta|\hat{\eta}) \\ & + (1 - F_1(\eta_2|\hat{\eta}))(1 - G(\Delta))\Delta. \end{aligned}$$

The training level unionised firms offer (in case they decide to train the worker), $\tilde{\tau}_u$, solves ¹⁰

$$\begin{aligned} c'(\tilde{\tau}_u) &= \frac{\partial E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}]}{\partial \tau} \\ &= \int_{\eta_1}^{\eta_2} (1 - G(y(\tilde{\tau}_u, \eta) - \bar{w})) \frac{\partial y}{\partial \tau} dF_1(\eta|\hat{\eta}) \\ &\quad - \int_{\eta_1}^{\eta_2} g(y(\tilde{\tau}_u, \eta) - \bar{w})(y(\tilde{\tau}_u, \eta) - \bar{w}) \frac{\partial y}{\partial \tau} dF_1(\eta|\hat{\eta}). \end{aligned} \tag{3.3}$$

Training affects profits in two ways. First, training increases the rent on trained workers. This effect is represented by the first term in (3.3). Second, training decreases

¹⁰Here, we have used that $y(\tau, \eta_2) = \bar{w} + \Delta$.

the probability that the worker stays with the firm. This effect is captured by the second term in (3.3). In appendix B we show the first effect dominates the second effect. Hence, training increases the future profit of unionised firms, i.e. $\frac{\partial E[\Pi_u \hat{\eta}]}{\partial \tau} \geq 0$ and $\tilde{\tau}_u \geq 0$.

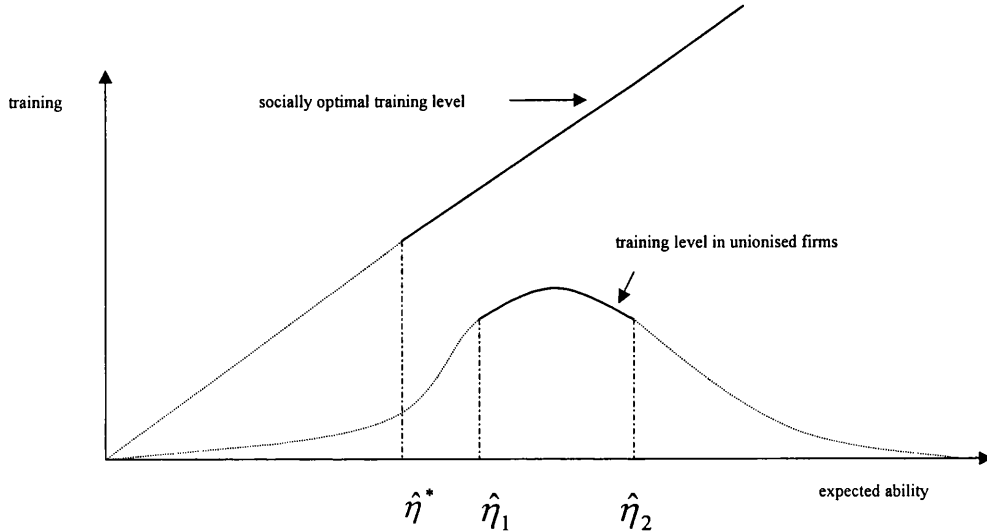
As the productivity of workers in training differs from the productivity of untrained workers, firms do not find it profitable to train every worker. The unionised firm will only offer training if the profit with training exceeds the profit without training. In appendix B we show that there exist two thresholds which we denote by $\hat{\eta}_1$ and $\hat{\eta}_2$. The unionised firm trains the worker if her expected ability lies in between these two thresholds. The training choice of unionised firms therefore satisfies

$$\tau_u^* = \begin{cases} 0 & \text{if } \hat{\eta} < \hat{\eta}_1 \text{ or } \hat{\eta} > \hat{\eta}_2, \\ \tilde{\tau}_u & \text{if } \hat{\eta}_1 \leq \hat{\eta} \leq \hat{\eta}_2. \end{cases}$$

The intuition for these results is as follows. Recall from figure 3.1 that training increases future profits only of workers with ability between η_1^t and η_2^{nt} . Consider a worker with very low expected ability. This worker is likely to turn out to be less able than η_1^t . The probability that she will be laid off after training is therefore very high, and the firm is better off by not training (and not hiring) her. A worker with a very high expected productivity, on the other hand, is likely to turn out to be more able than η_2^{nt} . The probability that this worker will be affected by the union wage is therefore low even without training. Again, training has only a small impact on the firm's future profit. In other words, union agreements compress wages only for workers with a productivity around the union wage, and firms find it most profitable to train these workers. Consequently, the impact of union agreements on training is not uniform: Unions have little or no impact on training for workers with very low and very high (expected) productivity, and the strongest impact for workers with expected productivity around the union wage.

For the US, there is evidence that more productive workers receive more training (e.g. Altonji and Spetzler 1991). This should not be interpreted as evidence against our model. There are other reasons for firm-financed general training, such as the complementarity between firm-specific and general training, which we have not modelled

Figure 3.2: Socially optimal training and training in unionised firms



and which may lead to more training for more able workers. Our model merely predicts that the impact of unions is strongest for workers in the middle range of the ability distribution.

Figure 3.2 plots the socially optimal training level and the training choice of unionised firms as a function of workers' expected ability. Non-unionised firms train no worker. Unionised firms offer some training. It is of interest to compare the training level unionised firms offer with the socially optimal training level. In the social optimum, the marginal cost of training is equal to the marginal product of training. Since there is a fixed cost of training, only workers with an expected ability above $\hat{\eta}^*$ are trained (see appendix A for details). Since unionised firms choose training such that its profits are maximised, and ignores the impact of training on the utility of the worker, unionised firms offer less training the socially optimal.

Next, we analyse wage determination in unionised and non-unionised firms in the first period. Details can be found in appendix B. Wages in the first period are determined by the firm's zero profit condition. As firms make positive profits in the second period, they make negative profits in the first period. Since union agreements do not affect profits and wage determination in non-unionised firms, non-unionised firms offer

the same first period wage as in the absence of union agreements, (3.2). The first period wage offer of unionised firms can be similarly derived. It is the firm that bears the training cost. We summarise our results in the following proposition.

Proposition 2 *Under limited commitment non-unionised firms offer no training. Unionised firms train workers with expected ability $\hat{\eta}_1 < \hat{\eta} < \hat{\eta}_2$. These workers are offered a training level of $\tilde{\tau}_u$. Training in unionised firms is less than socially optimal. Firms bear the training cost.*

Proof. See appendix B. ■

Sorting into the unionised sector

We finally sketch the sorting of workers into the unionised and non-unionised sector in the first period. Details can be found in appendix C. A worker bases her decision in which sector to work on the training level as well as on the first and second period wage unionised and non-unionised firms offer. She chooses to work in the sector in which her utility is higher.

Consider first the impact of training on worker sorting. As it is the firm which bears the training cost, workers are better off if they receive training. Hence, workers for whom union agreements increase the probability of receiving training ($\hat{\eta}_1 < \hat{\eta} < \hat{\eta}_2$) typically prefer to work in the unionised sector. *Workers sort into the unionised sector based on the impact unions have on training.*

Second, consider the impact of the worker's expected ability on worker sorting. Suppose unionised and non-unionised firms offer the same training level. Then there exist an ability threshold such that workers with expected ability above this threshold prefer to work in unionised firms. *More able workers thus self-select into the unionised sector.* The intuition for this result is simple. *Ex post*, workers who will be paid the union wage are better off, while workers who turn out to be less productive than the union wage are worse off when working in the unionised sector. Workers with low expected ability are likely to be of lower productivity than the union wage, and thus choose to work in non-unionised firms. We would like to emphasise that *ex ante* all workers are better off due to unions. Since a worker can always choose to work in

a non-unionised firm, she is guaranteed at least the same payoff as in the absence of union agreements. This is essentially due to our assumption that there are no spill-overs from unionised to non-unionised firms: Wage determination in non-unionised firms is unaffected by unionised firms.

Also note that in our model firms are indifferent between becoming unionised or not. Both unionised and non-unionised firms make zero profits in the long-run. Compared to non-unionised firms unionised firms make a lower profit in the second period and therefore a lower loss in the first period. There is thus less front-loading in unionised firms.

Interpretation: Unions as a commitment device

We next give an interpretation of why unions increase training in the economy. Recall that the reason why -in the absence of unions- the training market breaks down is that firms cannot fully commit to training provision. This problem can in principle be mitigated -though not eliminated- by a long-term wage contract that does not only specify today's but also future wages. In our model, unionised firms offer a special type of a long-term wage contract: They guarantee to pay at least the union wage in the future. In principle, firms could offer such a contract without becoming unionised. The problem with this type of contract is that it is not self-enforceable. Once training is completed, the firm has an incentive to deviate and pay a lower wage than the agreed minimum wage. Hence, the role unions essentially play in our model is that they serve as a commitment device. Unionised firms credibly signal to workers that they will pay at least the agreed union wage in the future. This then provides an incentive for firms to train workers, and improves welfare in the economy.

3.3 Empirical Implications

3.3.1 Training in unionised and non-unionised firms

According to our model, union agreements compress wages only in unionised, but not in non-unionised firms, therefore predicting training in unionised, but not in non-

unionised firms. Our model addresses only one reason for wage compression - other reasons include asymmetric information. This may lead to firm-financed training also in non-unionised firms. The key test for our model is whether unionised firms are more likely to train workers in apprenticeship schemes than non-unionised firms.

Testing this hypothesis is not straightforward for a number of reasons. First, our model suggests that workers with a higher expected ability self-select into the unionised sector. If more able workers are more likely to receive training, a simple comparison of the mean training intensity in unionised and non-unionised firms will overstate the causal impact of unions on training. Second, our model suggests that the impact of unions on training depends on workers' expected ability, and that workers sort into the unionised sector based on the impact unions have on training. This may lead to an upward biased estimate for the *average* impact unions have on training (i.e. the impact of unions on the training probability of a randomly selected worker or firm). Finally, the firm's decision to be unionised may also depend on firm characteristics. If these characteristics are correlated with the firm's propensity to train, a simple comparison of mean training intensity is again misleading.

Our empirical test is based on firm panel data, supplemented by information about the firm's workforce. We use a difference in difference estimator to evaluate the effects of unionisation on training, assuming that changes over time are the same in firms that change unionisation status, and in firms that remain non-unionised in both periods. To take account of changes in firm characteristics as well as workforce quality within the firm, we condition on changes in observable worker and firm characteristics in addition.

Define \bar{y}_{jt} as the proportion of apprentices at time t in firm j . The proportion of workers in apprenticeship schemes depends on the union status of the firm U_{jt} , a time effect θ_t , observed and unobserved average worker characteristics, $\bar{\eta}_{jt}$, as well as observed and unobserved firm characteristics, f_j . Both $\bar{\eta}_{jt}$ and f_j are defined as deviations from the population mean. Assuming linearity, this relationship can be written as

$$\bar{y}_{jt} = \beta + \lambda_{jt} U_{jt} + \theta_t + \bar{\eta}_{jt} + f_j + v_{jt}, \quad (3.4)$$

where v_{jt} is an i.i.d. error term, and λ_{jt} is the effect of unions on the training probability. This parameter is time and firm-specific, as the impact of unions on training depends on workers' expected ability. The parameter we seek to identify is the difference in the training intensity between being unionised and non-unionised for those firms that choose to be unionised, $E(\bar{y}_{jt}^1 - \bar{y}_{jt}^0 | U_{jt} = 1)$. Evaluation of this effect requires to construct a missing counterfactual, namely the training intensity of firms that choose to be unionised, in the non-unionised state.

To identify this effect, we make use of the panel nature of our sample. We observe the training intensity of firms that enter union agreements in period t in previous periods when they were not unionised. However, a simple difference estimator may confound the causal effect of unionisation on training with common time effects, and changes in the composition of the workforce, as -according to our model better workers self-select into unionised firms.

We assume that changes in common time effects ($\Delta\theta_t$) are the same in firms that are not unionised in both time periods, and in firms that change from being non-unionised to being unionised, conditional on changes in observables. Furthermore, we assume that any variation in changes in the workforce quality that is correlated with changes in the union status is absorbed by changes in observed worker characteristics as well as firm characteristics. These assumptions imply that

$$\begin{aligned} m_{jt} &= E(\Delta\theta_t + \Delta\bar{\eta}_{jt} + \Delta v_{jt} | U_{jt} = 0, U_{jt-1} = 0, \Delta X_{jt}) \\ &= E(\Delta\theta_t + \Delta\bar{\eta}_{jt} + \Delta v_{jt} | U_{jt} = 1, U_{jt-1} = 0, \Delta X_{jt}), \end{aligned}$$

where X_{jt} is a vector of observed characteristics of the firm's workforce as well as observed firm characteristics. The key identifying assumption is that, conditional on changes in X_{jt} , changes in the counterfactual outcome distribution of firms who changed to be unionised, had they not changed to be unionised, is the same than the observed outcome distribution of firms that are not unionised. Under this assumption, estimation identifies the treatment on the treated effect:

$$E(\bar{y}_{jt}^1 - \bar{y}_{jt}^0 | U_{jt} = 1) = E(\bar{y}_{jt} - \bar{y}_{jt-1} | U_{jt} = 1, U_{jt-1} = 0, \Delta X_{jt}) - m_{jt}.$$

Our estimation strategy combines a linear matching estimator with a difference in difference estimator. The equation we estimate is given by

$$\Delta \bar{y}_{jt} = \Delta \theta_t + \Delta X'_{jt}d + \Delta v_{jt} + \lambda_{jt}U_{jt}, \quad (3.5)$$

where the sample includes firms that are non-unionised in both periods, and firms that change from being non-unionised in period $t - 1$ to being unionised in period t .

Another parameter of interest is the effect of leaving the unionised sector on training intensity. Under the same set of assumptions than above, this should then identify $E(\bar{y}_{jt}^0 - \bar{y}_{jt}^1 | U_{jt} = 0)$. Furthermore, we could define as an alternative comparison group firms that are unionised in both periods. In this case, identification requires the additional assumption that $E(\lambda_{j(t)} - \lambda_{j(t-1)} | U_{jt} = 1, U_{j(t-1)} = 1, \Delta X_{j(t)}) = 0$. Below we present results for all these groups.

3.3.2 Worker selection: Wages of movers and stayers in unionised and non-unionised firms

The second set of empirical implications of our model concerns worker selection. These additional implications allow us to test whether the assumptions we made about wage determination in unionised and non-unionised firms are compatible with the data. They also help us to identify other reasons for firm-financed training, and thus help to answer the question why possibly also non-unionised firms offer apprenticeship training. However, they hold even if unions have no impact on training.

In our model, two types of worker selection are present. First, in period one workers sort into the unionised and non-unionised sector based on their expected ability. Second, in the second period, workers' decision to leave the training firm depends on their revealed ability.

The following proposition summarises the empirical implications of our model for this second type of selection. In the empirical analysis, we control for the first type of selection.

Proposition 3 (i) *In non-unionised firms, endogenous and exogenous movers are as productive as stayers and earn a higher wage than stayers.* (ii) *In unionised firms,*

*endogenous and exogenous movers **may** be less productive and **may** earn a lower wage than stayers. (iii) The difference between the wage of endogenous movers and stayers is lower in unionised than in non-unionised firms.*

Proof. See appendix D. ■

The intuition for these results is as follows. Consider first workers in non-unionised firms. Perfect competition ensures that workers who leave the training firm are paid their marginal product, while workers who remain with the incumbent firm receive a wage that is lower than their marginal product, due to a rent Δ induced by non-pecuniary utility shocks. Hence, conditional on productivity, endogenous as well as exogenous movers earn a higher wage than stayers. Furthermore, the difference between the incumbent and outside offer, Δ , does not depend on workers' productivity. Consequently, the probability that a worker leaves the firm does not depend on her productivity, and stayers, endogenous as well as exogenous movers all have the same productivity on average.

Next, consider workers in unionised firms. As in non-unionised firms, exogenous and endogenous movers earn a higher wage than stayers conditional on productivity. However, in unionised firms there is also selection. On the one hand, unionised firms dismiss all workers with a productivity below the union wage after training. These workers find employment in non-unionised firms and are paid a wage equal to their productivity. This effect leads to a lower average productivity of movers than of stayers. On the other hand, among workers who are paid the union wage, it is the more able workers who are more likely to leave. Figure 3.1 illustrates this point. Recall that the probability of moving is $G(v - w)$, where w and v denote the worker's training and outside firm wage offer, respectively. Consider a trained worker with ability η_1^t . This worker is offered the same wage by the training and outside firms. Hence, the probability that she leaves the training firm is $G(0)$. Workers with ability above η_1^t , on the other hand, are offered a higher wage by outside firms than by the training firm, and are thus more likely to leave the training firm. This effect leads to a higher average productivity of movers than of stayers.

Finally, the model implies that the difference between the wage of endogenous

movers and stayers is lower in unionised than in non-unionised firms. The reason is that for workers who are paid the union wage -and among whom it is the more able workers who are more likely to leave the training firm-, the difference between the incumbent and outside wage offer is less than Δ , i.e. what it would be in non-unionised firms.

Unconditional on the union status our model thus makes no ambiguous predictions about wages of movers and stayers. Hence, our model's empirical implications are consistent with empirical predictions of Acemoglu and Pischke's (1998) asymmetric information model: Workers who leave the training firm for endogenous reasons are less able than workers who stay with the training firm or leave the training firm for exogenous reasons (i.e. because of plant closure or military service). Information on the union status of the training firm helps to distinguish between these two alternative explanations, as our model unambiguously predicts a higher mover-stayer wage differential in unionised than in non-unionised firms.

Our empirical test is based on the following simple model¹¹:

$$w_{it} = \gamma M_{i(0)} + \bar{\eta}_{j(0)} + (\eta_i - \bar{\eta}_{j(0)}) + e_{ijt},$$

where w_{it} denotes the worker i 's log wages t periods after apprenticeship completion. $M_{i(0)}$ is an indicator variable, being equal to one if the worker left the training firm right after the training period. We will distinguish between separation for exogenous or endogenous reasons. Endogenous movers are workers who leave the training firm either because they are laid off, or because they quit. Exogenous movers leave the training firm because for exogenous reasons (we use military service and training firm closure in our analysis).

Separation may be separation from a unionised firm ($M_{i(0)}^u$), or a separation from a non-unionised firm ($M_{i(0)}^{nu}$), with $M_{i(0)} = M_{i(0)}^u + M_{i(0)}^{nu}$. The quality of the worker is represented by η_i , while $\bar{\eta}_{j(0)}$ captures the average quality of workers at firm j at the time of apprenticeship completion that trained the worker. When we say worker quality, we mean both her ability and training she received. Finally, e_{ijt} is a random

¹¹See von Wachter (2002) for a similar set-up.

disturbance term. Due to non-pecuniary job characteristics, our union model implies $\gamma > 0$, for workers trained in unionised as well as non-unionised firms. OLS estimation results in

$$\hat{\gamma} = \gamma + \frac{\text{Cov}(\bar{\eta}_{j(0)}, M_{i(0)})}{\text{Var}(M_{i(0)})} + \frac{\text{Cov}(\eta_i - \bar{\eta}_{j(0)}, M_{i(0)})}{\text{Var}(M_{i(0)})}. \quad (3.6)$$

The second term in (3.6) captures sorting of workers into training firms. In our union model, workers sort into the unionised sector in the first period based on their expected ability. If retention rates after apprenticeship completion are higher in unionised than in non-unionised firms, then the second term is negative, otherwise it is positive.

The third term captures the selection of workers after training. It tells us whether among trainees in the same firm, it is the less or more able workers who are more likely to leave. It is this covariance we are interested in. Acemoglu and Pischke's (1999) asymmetric information model implies a negative covariance, while our union model makes no unambiguous prediction. Suppose we can control for sorting of workers into training firms. Then, as $\gamma > 0$, a negative estimate for $\hat{\gamma}$ -i.e. a lower wage for movers than for stayers- implies that it is the low quality workers who are more likely to leave the training firm. This is what we call adverse selection. In this case, the wage difference between movers and stayers should be permanent, and visible even many years after apprenticeship completion.

We control for selection into training firms in two ways. First, we condition on observable training firm characteristics, such as firm size and sector. Second, there are some workers in our data who have done their apprenticeship in the same firm. This allows us to condition also on training firm fixed effects.

We also estimate this model for exogenous movers. A comparison between $\hat{\gamma}$ for exogenous and endogenous movers gives further evidence on adverse selection. Since both exogenous and endogenous movers are paid a wage equal to their marginal productivity, the difference between the estimates $\hat{\gamma}$'s identifies the productivity difference between the two groups of workers.

3.4 Empirical Analysis

3.4.1 Data Sources

Our empirical analysis makes use of three unique data sources for testing implications of our model. The first data source is a firm panel data set, covering the period between 1996 and 1999. To this data source, we have matched information of the firm's workforce, drawing on the population of workers from social security records. We use this data to compare the training intensity in unionised and non-unionised firms. The second data source is a one percent sample of the German workforce, drawn from social security records between 1975 and 1995. We use this data to test for an adverse selection of movers. This data does not contain information on the firm's union status. For 1995 and 1997, we are able to combine information on workers' post-training wages with information on the union status of the training firm. We use this data to compare the mover-stayer differential for workers trained in unionised and non-unionised firms. We describe the data sources and the samples we use for analysis in detail in appendix E.

3.4.2 Training in unionised and non-unionised firms

Table 3.1 displays information on union coverage, the proportion of trainees in unionised and non-unionised firms, and characteristics of firms and their workforce for the years 1996-1999. Information on the workforce (like the proportion of trainees, ratio of female employees, average age of workers, and proportion of qualified workers) are constructed by matching information of the workforce, retrieved from social security records, to the firm panel data, as described in the appendix E. Entries are weighted so that they are representative for firms (upper panel) and workers (lower panel).

About 56 percent of all firms are unionised. However, as unionised firms are predominantly large firms, 76 percent of all workers are employed in unionised firms. Overall, 29 percent of all firms train workers on apprenticeship schemes, but unionised firms are considerably more likely to train than non-unionised (38 percent vs. 17.5 percent). Those firms that are unionised have an average proportion of apprentices

Table 3.1: Unionised and non-unionised firms

	All		Unionised		Non-Unionised	
	56.20		Representative for firms			
Proportion Unionised	Mean	StdD	Mean	StdD	Mean	StdD
Proportion trainees (in percent)	8.09	17.98	10.18	19.17	5.42	15.93
Firm trains (in percent)	28.79	45.28	37.59	48.43	17.50	38.01
Proportion qualified workers	49.03	29.72	52.56	28.43	44.50	30.71
Number of new hires	0.999	6.78	1.16	8.17	0.783	4.38
Size	18.16	125.80	24.62	164.39	9.86	36.60
(Investment/worker)/1000*	18.77	603.39	11.81	322.09	27.66	835.05
(Turnover/worker)/1000*	383.42	1735.44	450.76	2086.94	294.61	1107.63
Average Age workers	37.22	8.36	36.68	8.06	37.92	8.69
Ratio females	23.84	63.86	24.26	31.21	23.30	89.78
Daily Average Wage	88.73	46.0	93.74	43.5	82.29	48.2
Proportion young firms (≤ 5)	20.67	40.49	19.22	39.40	22.52	41.78
Proportion old firms (≥ 30)	30.08	45.86	36.85	48.24	21.39	41.01
Profit evaluation good/very good	30.07	45.86	28.24	45.02	32.42	46.81
Profit evaluation bad	9.77	29.69	9.03	28.67	10.72	30.94
Representative for Workers						
Proportion working in unionised firms	76.19					
	Mean	StdD	Mean	StdD	Mean	StdD
Proportion trainees (in percent)	5.97	10.82	6.29	10.46	4.93	11.83
Firm trains (in percent)	59.65	46.06	67.74	46.74	33.75	47.29
Proportion qualified workers	58.15	29.39	60.5	28.06	50.46	32.09
Number of new hires	25.23	105.58	30.3	118.84	9.02	35.86
Size	889.48	3431.66	1121.8	3889.27	145.60	570.68
(Investment/worker)/1000*	16.12	385.82	14.6	240.10	20.71	659.46
(Turnover/worker)/1000*	711.14	5406.14	829.3	6133.22	320.32	1141.48
Average Age workers	38.52	5.52	38.6	5.07	38.16	6.77
Ratio females	21.83	53.46	21.4	24.05	23.00	100.78
Daily Average Wage	118.16	50.37	123.89	48.32	99.80	52.36
Proportion young firms (≤ 5)	15.61	36.30	13.3	33.98	22.94	42.05
Proportion old firms (≥ 30)	52.58	49.93	60.4	48.88	27.25	44.53
Profit evaluation good/very good	31.00	46.25	29.1	45.45	36.91	48.26
Profit evaluation bad	9.94	29.92	10.4	30.56	8.39	27.73

Firm data and matched worker characteristics, 1996-1999. *: In 1996 German Marks

of 10 percent, while those firms that are not unionised have a proportion of only 5.42 percent. The numbers in the lower panel indicate that about 60 percent of workers work in firms that report that they train apprentices; 34 percent of workers working in non-unionised firms work in firms that train, while this is the case for 68 percent of workers employed in unionised firms. These numbers show that also non-unionised firms are engaged in apprenticeship training. However, in accordance with our model,

the training intensity is higher in unionised firms.

The remaining table entries compare unionised and non-unionised firms in terms of observable characteristics. The numbers show that unionised firms are substantially larger and older than non-unionised firms. There is some evidence that the turnover per worker is higher in unionised firms, while investment per worker and firms' evaluation of profits is more favourable in non-unionised firms.

In table 3.2, we display results from a difference in difference - matching estimator (see (3.5) in section 3.3.1). The set of matching variables includes firm characteristics, and characteristics of the workforce, as explained in the footnote of the table. The first pair of columns uses firms that change from being non-unionised to being unionised, where the reference category are firms that are non-unionised over the entire sample period. Under the assumptions outlined in section 3.3.1, these estimates identify the difference in the proportion of trainees (or in the propensity to train, row 2 of the table) between being unionised and non-unionised for unionised firms. The estimates show that the effect of unionisation on the proportion of apprentices is positive; the point estimate is about 2.7 percent. Results in the second row refer to training propensity; again, the estimate of the effect is positive and significant, indicating a 6.8 percentage points higher probability of training in unionised firms.

The results in the next pair of columns refers to firms that change from being unionised to being non-unionised, where again non-unionised firms identify the time effects. The effect on the proportion of apprentices is practically zero, but not negative, as predicted by our model. A reason may be that downward adjustment is not immediate. Apprenticeship training schemes last for 2 - 3 years, and reduction is therefore only possible at the margin. The effect on the self reported information of whether the firm trains is negative, but insignificant.

In the next two pairs of columns, we repeat the same regressions, but we now identify the time effects from firms that are unionised over the course of the panel. As we explain above, this implies the additional assumption that the selection into unionised firms over time remains constant. The results are very similar to those in columns 1 and 2, and are further support for our model.

Table 3.2: Effect of Unionisation on Training Intensity

	Reference: Nonunion-Nonunion				Reference: Union-Union			
	Nonunion-Union	Union-Non-Union	Nonunion-Union	Union-Non-Union	Nonunion-Union	Union-Non-Union	Nonunion-Union	Union-Non-Union
Proportion Apprentices	0.0267	0.0104	0.0071	0.0059	0.0306	0.0067	0.0032	0.0041
Firm Trains	0.0677	0.0327	-0.0015	0.0192	0.0657	0.0297	0.0020	0.0181
	1100		1513		4478		4891	

Source: Firm panel, matched with employee information, 1996-1999. All regressions include time dummies. Firm characteristics include changes in firm size, investment per worker, turnover per worker, number of new hires, and evaluation of profit. Worker characteristics include changes in the proportion of qualified workers, the average age of workers, the average daily wage, and the ratio of females.

3.4.3 Worker selection

We now investigate the second set of empirical implications of our model, which concerns worker selection. These additional implications allow us to test whether the assumptions we made about wage determination in unionised and non-unionised firms are compatible with the data. They also help us to identify possible other reasons for firm-financed training, such as asymmetric information, and thus help to answer the question why also non-unionised firms offer apprenticeship training.

We commence by estimating the wage differential for workers who stay with, and leave the apprenticeship firm, using longitudinal data. This data does not contain information on the worker's union status. It has, however, several other advantages. First, it allows us to precisely distinguish between stayers, exogenous and endogenous movers. Second, and most importantly, we can study the evolution of wages, thus analyse whether differences between wages for movers and stayers are permanent. Third, for large firms, we are able to condition on training firm effects, thus eliminating the effects of worker selection into training firms.

We then use matched employee-employer data, and distinguish between workers trained in unionised and non-unionised firms. This data has the disadvantage that we observe workers' wages only once, right after apprenticeship completion. Further-

more, we cannot distinguish between exogenous and endogenous movers, and it does not constitute a random sample of apprentices. Despite these drawbacks, it helps to discriminate between our union model and an asymmetric information model. Our union model predicts a higher wage for movers than for stayers in non-unionised firms. Hence, a higher wage for stayers among workers trained in non-unionised firms must be due to asymmetric information. A higher mover-stayer wage differential in unionised firms is evidence that union agreements contribute to an adverse selection of movers.

In table 3.3, we report wage differentials between stayers and various groups of movers conditional on individual characteristics (columns 1) and individual as well as training firm characteristics (columns 2). The entries in the two columns are coefficients on the respective dummy variable (for exogenous and endogenous movers).

Mobility after apprenticeship training is considerable. Overall, 33 percent of workers leave the training firm after training. Of those, 6.9 percent are forced to leave their training firm due to closure. If we do not condition on firm characteristics, workers who remain with the training firm earn a 10.23 percent higher wage than workers who leave the training firm for endogenous reasons. This wage differential is highly significant. It reduces to 5.68 percent when we condition on training firm characteristics in addition. This suggests that there is sorting of workers into training firms.

A comparison between wages for exogenous and endogenous movers (row 2 and 3) show sensitivity to the set of conditioning variables. When conditioning only on worker characteristics, exogenous movers earn a significantly lower wage than endogenous movers. However, it is predominantly small firms which close down¹². If larger firms attract better workers, and larger firms have higher retention rates, then this could explain the difference. Indeed, when we condition on characteristics of the training firm, exogenous movers obtain wages that are 3.3% higher than endogenous movers. This difference is significant at the 1 percent level. Both groups obtain lower wages than individuals who remain with their training firm after the training period has finished.

¹²The average size (median) of the apprenticeship firm for stayers is 1560 (57), for endogenous movers 734 (35), for workers who leave due to plant closure 130 (8), and for workers who are drafted into military service 560 (34).

Table 3.3: Wages of stayers, endogenous and exogenous movers

			1		2	
			worker	standard	+ training firm	standard
			characteristics	errors	characteristics	errors
			N=24194		N=24194	
1	stays	67 %				
2	leaves	30.7 %	-0.1023	(0.0041)	-0.0568	(0.0039)
3	leaves, plant closure	2.3 %	-0.1394	(0.0128)	-0.0249	(0.0119)
	Test: 2=3		F(1,24173)=8.14	Pv: 0.004	F(1,24162)=7.35	Pv: 0.006
4	no service, stays	62.8 %				
5	service, returns	3.6 %	0.0005	(0.0100)	0.0004	(0.0099)
6	service, leaves	7.7 %	-0.0878	(0.0076)	-0.0409	(0.0070)
7	no service, leaves	25.9 %	-0.1057	(0.0046)	-0.0582	(0.0043)
	Test: 6=7		F(1,22174)=4.72	Pv: 0.029	F(1,22163)=5.4	Pv: 0.021

Standard errors in parenthesis. The first column controls for worker characteristics only. Regressors included are age at the end of the apprenticeship, duration of the apprenticeship, A-levels (Abitur), nationality and time effects. Column 2 also controls for characteristics of the apprenticeship firm, i.e. firm size (in logs) and sector. For regressions that include service quitters, "problematic cases" are dropped. This reduces the sample size to 22196. Definition 1 is used for service quitters. Results are not sensitive to the definition used and whether problematic cases are dropped or not.

The next panel of table 3.3 (rows 4-7) breaks the sample down according to whether individuals went to the service after their apprenticeship training. This is the case for 11.3 percent of our sample. Of those, 68 percent do not return to the training firm. Those who do service and return to the training firm have virtually identical wages to those who stay with the training firm (row 6). In consequence, any wage advantage of workers who do not return to the training firm after their service, compared to endogenous movers, is unlikely to be due to training received whilst in service¹³. Those who were drafted into the service right after the apprenticeship and did not return to their training firm have higher wages than endogenous movers (rows 7 and 8). This difference is significant at a 5 percent level for both specifications.

¹³This confirms Acemoglu and Pischke's (1998) conjecture that military service does not increase productivity.

These results are in line with an adverse selection of movers. However, our observable training firm characteristics may be insufficient to control for sorting of workers into training firms. Next, we present results that condition on fixed training firm effects.

Fixed training firm effects

In our sample, we observe for some larger establishments more than one worker who obtains his apprenticeship training. This sample selects on large firms - while the average (median) firm size is 1265 (58), the average (median) size of firms with multiple apprentices is 5602 (1264). In table (3.4), we present results. Entries in the first and second row replicate OLS estimations conditional on worker (row 1) and worker as well as training firm characteristics (row 2); thus, results correspond to those reported in the first and second column in table 3.3. For this selected sample of workers, OLS estimates are 4-5 percentage points larger than those using the overall sample.

Table 3.4: Mover, fixed training firm effect estimates

	Coeff	StdE
1 leaves, worker characteristics	-0.1437	(0.0092)
2 leaves, worker+firm characteristics	-0.1079	(0.0096)
3 fixed training firm effect	-0.1059	(0.0095)

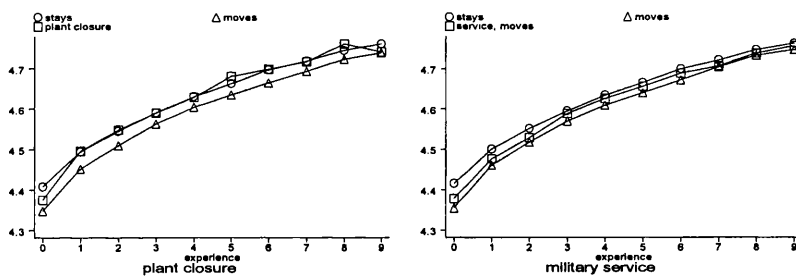
N=3874. Same specification as table 3.3.

The third row reports estimates conditional on training firm fixed effects. The estimate is virtually identical to the one that conditions on observable training firm characteristics. This suggests that there is little sorting into training firms, conditional on training firm size and sector. The lower wage for endogenous movers is therefore likely to be a consequence of adverse selection. If this is the case, the wage difference between movers and stayers should be persistent, and visible many years after apprenticeship completion. We next analyse the evolution of the mover-stayer wage differential by experience.

Are differentials persistent?

In figure 3.3, we plot wages against experience, where we distinguish between stayers, movers, and individuals who change firm because the training firm closes down (left panel) and individuals who change firms because of service (right panel). The graphs are based on regressions which include the same set of variables as columns 2 of table 3.3. The graphs are constructed from estimated coefficients for experience dummies, and experience dummies interacted with an indicator variable for endogenous move, or the respective exogenous move.

Figure 3.3: Wages of stayers, endogenous and exogenous movers by experience



Wage disadvantages of endogenous movers relative to stayers are persistent, and remain sizable even after 10 years of labour market experience. Both groups of exogenous movers have initially an intermediate position, but their wages catch up with those of stayers. This provides additional support for the hypothesis that less productive workers are more likely to leave the training firm.

We have also estimated regressions that exclude sectors commonly viewed as problematic, like agriculture, and construction¹⁴. Results change only slightly compared to those reported above.

To sum up, we find strong evidence for an adverse selection of movers: It is the less able trainees who are more likely to leave the training firm. This is implied by Acemoglu and Pischke's asymmetric information model. It is, however, also consistent with our union model. In the next section, we use matched data to discriminate between these

¹⁴Acemoglu and Pischke (1998) report that in the construction sector a penalty is imposed on firms who do not train apprentices.

two reasons for an adverse selection of movers.

Matched data: Unionised versus non-unionised firms

We now break down the mover-stayer wage differential by the union status of the worker's training firm. Our union model predicts a higher wage for movers than for stayers in non-unionised firms. Hence, a higher wage of stayers among workers trained in non-unionised firms must be due to asymmetric information. A higher mover-stayer wage differential in unionised firms is evidence that union agreements contribute to an adverse selection of movers.

We run separate regressions for workers trained in unionised and non-unionised firms. Regressions include the same set of variables as columns 2 of table 3.3: dummy variables for A-levels, nationality, year, and sector of apprenticeship firm; furthermore, the duration of apprenticeship, age, and the firm size of the apprenticeship firm.

Table 3.5: Wages of movers and stayers: Unionised versus non-unionised firms

	Unionised N=35067			Non-unionised N=692		
	proportion	coefficient	StdE	proportion	coefficient	StdE
stays	74.34 %			68.83 %		
leaves	25.66 %	-0.1481	(0.0042)	31.17 %	-0.0771	(0.046)

Standard errors in parenthesis. Regressors included are age at the end of the apprenticeship, duration of apprenticeship, A-levels (Abitur), nationality, time effects, firm size (in logs) and sector of training firm.

Table 3.5 reports results. We find a higher wage for stayers than movers also for workers trained in non-unionised firms, indicating that asymmetric information is one reason for an adverse selection of movers. However, the wage differential is 7.1 percentage points larger for workers trained in unionised firms, suggesting that union agreements contribute to an adverse selection.

3.5 Discussion and Conclusion

In this paper we revisit the question how wage rigidities caused by unions affect firm-financed training. We address this questions within the German apprenticeship system. We develop a simple model of unions and firm-financed training, and test its empirical implications using matched firm-worker data. Our model is intended to capture the specific features of the German collective bargaining system. In Germany, only firms that belong to an employer federation are legally obliged to pay at least union wages. Membership in an employer federation is voluntary. Union wages act as minimum wages and apply to all workers in a unionised firm. The most important difference between our and existing models, such as Acemoglu and Pischke (1999b) and Booth et al. (1999), therefore is that in our model unionised and non-unionised firms coexist. Unionised firms have to pay at least the union wage, but are allowed to pay a higher wage¹⁵. Non-unionised firms, in contrast, may pay whatever wage they want. Workers differ with respect to ability, but the same union wage applies to all of them. Within this framework, we derive wage and training determination in unionised and non-unionised firms, as well as the sorting of workers into unionised and non-unionised firms. Workers choose to work in the sector in which their utility is higher. Firms are indifferent between joining the unionised or non-unionised sector.

Our key theoretical results can be summarised as follows. First, an adverse selection of movers is not unique to Acemoglu and Pischke's asymmetric learning model, but is also consistent with our union model. Second, union agreements compress wages for workers with a productivity around the union wage in *unionised* firms, inducing them to train these workers. If union agreements are the only source of wage compression, non-unionised firms offer no training. Training is less than socially optimal. Third, the impact of unions on training is not uniform. Workers sort into the unionised sector based on their ability and the impact union agreements have on training. The role unions essentially play in our model is that they serve as a commitment device. Unionised firms credibly signal to workers that they will pay at least the agreed union

¹⁵Our model therefore differs from union models such as Acemoglu and Pischke (1999b) and Booth et. al. (1999) who assume that firms are forced to pay the union wage.

wage in the future. Unions therefore make a simple form of a long-term wage contract enforceable.

Our empirical strategy proceeds in two steps. In a first step, we use firm panel data, matched with information about the firm's workforce, to directly test whether unions increase training. Our goal is to identify the causal impact of unions on training. Our estimation strategy follows a difference in difference - linear matching estimator, taking account of nonrandom selection of workers and possibly firms into the unionised sector. We consistently find that unions increase training.

In a second step, we test for an adverse selection of movers. We find strong evidence for this: Even after ten years in the labour market, endogenous movers earn considerably lower wages than stayers and exogenous movers. We then use matched employee-employer data to compare the mover-stayer wage differential for workers trained in unionised and non-unionised firms. We find a higher wage for stayers than for movers not only for workers trained in unionised firms but also for those trained in non-unionised firms, but the differential is twice as high for workers trained in unionised firms. This suggests that wage floors caused by unions contribute to an adverse selection of movers, but are not the only reason. We conclude that wage floors created by unions are an important reason for firm-financed training in Germany. However, we cannot rule out asymmetric information as an additional reason for firm-financed training.

Our results have important policy implications. First, a deregulation of the German labour market that limits the power of unions may have undesired consequences for the German apprenticeship system. Second, attempts to introduce training schemes in countries in which unions play only a minor role may fail. It is, however, important to bear in mind that the reason why unions induce firms to sponsor training is that unions make long-term wage contracts enforceable. There may be ways to do so other than through the unionisation of the economy -i.e. by strengthening the legal system. Furthermore, training is less than socially optimal since firms cannot fully commit to training provision. A better outcome may therefore be achieved by monitoring firms to insure commitment to training provision.

Our results also have important welfare implications. Most importantly, we find no evidence that unions reduce training in the economy. Moreover, our results are compatible with unions moving training closer to the socially optimal level. The welfare-improving impact of unions is the greater the higher the profit firms make. This suggests that in labour markets with low mobility costs and wages close to productivity -as maybe in the US- wage floors have only a small impact on training.

3.6 Appendix

A Socially optimal training

At the social optimum the marginal cost of training is equal to the marginal product of training:

$$\int_{-\infty}^{\infty} \frac{\partial y(\tilde{\tau}, \eta)}{\partial \tau} dF_1(\eta|\hat{\eta}) = c'(\tilde{\tau}).$$

The assumptions on the cost and production function ensure that the second order condition is satisfied. As the productivity of workers during training is smaller than the productivity of untrained workers (fixed cost of training), it is not optimal to train every worker. A worker should only be trained if her productivity when trained exceeds her productivity when not trained, i.e. if

$$\int_{-\infty}^{\infty} y(\tilde{\tau}, \eta) dF_1(\eta|\hat{\eta}) - c(\tilde{\tau}) - k - \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}) \geq 0. \quad (3.7)$$

We next show that only workers with expected ability above $\hat{\eta}^*$ are trained. Totally differentiating the left hand side of (3.2.1) with respect to $\hat{\eta}$ yields

$$\int_{-\infty}^{\infty} \frac{\partial y(\tilde{\tau}, \eta)}{\partial \eta} dF_1(\eta|\hat{\eta}) - \int_{-\infty}^{\infty} \frac{\partial y(0, \eta)}{\partial \eta} dF_1(\eta|\hat{\eta}) > 0$$

due to the complementarity between ability and training. There thus exists an ability threshold $\hat{\eta}^*$ such that all workers with expected ability greater than $\hat{\eta}^*$ are trained, where $\hat{\eta}^*$ is implicitly defined as

$$\int_{-\infty}^{\infty} y(\tilde{\tau}, \eta) dF_1(\eta|\hat{\eta}^*) - \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}^*) = c(\tilde{\tau}) + k.$$

The socially optimal training level τ^* therefore satisfies

$$\tau^* = \begin{cases} 0 & \text{if } \hat{\eta} < \hat{\eta}^*, \\ \tilde{\tau} & \text{if } \hat{\eta} \geq \hat{\eta}^*. \end{cases}$$

B Proof of proposition 2

Proposition 2 Under limited commitment non-unionised firms offer no training. Unionised firms train workers with expected ability $\hat{\eta}_1 < \hat{\eta} < \hat{\eta}_2$. These workers are offered a training

level of $\tilde{\tau}_u$. Training in unionised firms is less than socially optimal. Firms bear the training cost.

We first show that the future profit of the unionised firm, $E[\Pi_u(\tau, \eta)|\hat{\eta}]$, is increasing in training. The increase in $E[\Pi_u(\tau, \eta)|\hat{\eta}]$ due to training equals

$$\begin{aligned} \frac{\partial E[\Pi_u(\tau, \eta)|\hat{\eta}]}{\partial \tau} &= \int_{\eta_1}^{\eta_2} (1 - G(y(\tau, \eta) - \bar{w})) \frac{\partial y}{\partial \tau} dF_1(\eta|\hat{\eta}) \\ &\quad - \int_{\eta_1}^{\eta_2} g(y(\tau, \eta) - \bar{w})(y(\tau, \eta) - \bar{w}) \frac{\partial y}{\partial \tau} dF_1(\eta|\hat{\eta}). \end{aligned}$$

Recall that for workers with expected ability between η_1 and η_2 the union wage is greater than the wage the firm would choose optimally. Hence, from the first order condition of the second period wage, $1 - G(y - \bar{w}) \geq g(y - \bar{w})(y - \bar{w})$. Consequently, $\frac{\partial E[\Pi_u(\tau, \eta)|\hat{\eta}]}{\partial \tau} \geq 0$, and $\tilde{\tau}_u \geq 0$.

We next show that workers with expected ability between $\hat{\eta}_1$ and $\hat{\eta}_2$ are trained. A unionised firm trains if profits with training exceeds profit without training, i.e. if

$$E[\Pi_u(0, \eta)|\hat{\eta}] < -c(\tilde{\tau}_u) - k + E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}].$$

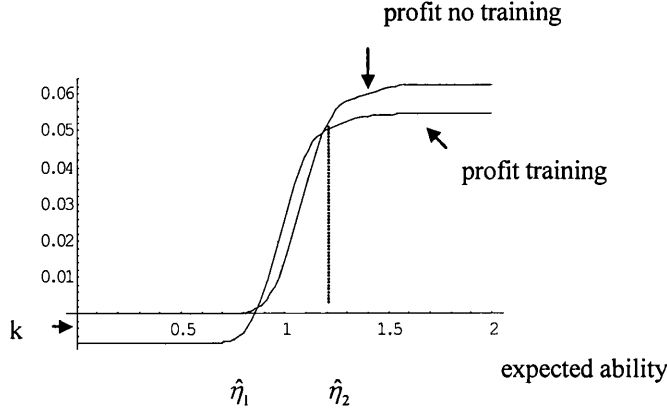
As $\frac{\partial E[\Pi_u(\tau, \eta)|\hat{\eta}]}{\partial \tau} \geq 0$, $E[\Pi_u(0, \eta)|\hat{\eta}] \leq -c(\tilde{\tau}_u) + E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}]$. Hence, if the fixed cost of training is equal to 0, ($k = 0$), every worker would get trained. Figure 3.4 illustrates the impact of the fixed cost of training on the probability that a worker receives training. It plots the firm's profit with and without training as a function of the worker's expected ability, for moderate (panel A) and high (panel B) values of the fixed cost of training, k . The firm's profit is increasing in the worker's expected ability, $\hat{\eta}$. It is first convex, then concave in $\hat{\eta}$. To prove this, we first show that $E[\Pi_u(\tau, \eta)|\hat{\eta}]$ is increasing in workers' expected ability. Differentiating $E[\Pi_u(\tau, \eta)|\hat{\eta}]$ with respect to $\hat{\eta}$ yields

$$\frac{\partial E[\Pi_u(\tau, \eta)|\hat{\eta}]}{\partial \hat{\eta}} = \int_{\eta_1}^{\eta_2} \{(1 - G(y - \bar{w})) - g(y - \bar{w})(y - \bar{w})\} \frac{\partial y}{\partial \eta} dF_1(\eta|\hat{\eta}) > 0.$$

We next show that Π_u is first convex and then concave in $\hat{\eta}$. Taking the second derivative yields

$$\frac{\partial^2 E[\Pi_u(\tau, \eta)|\hat{\eta}]}{\partial \hat{\eta}^2} = \int_{\eta_1}^{\eta_2} \{(1 - G(y - \bar{w})) - g(y - \bar{w})(y - \bar{w})\} \frac{\partial y}{\partial \eta} \frac{\partial f_1(\eta|\hat{\eta})}{\partial \hat{\eta}} d\eta.$$

Figure 3.4: Profit with and without training as a function of workers expected ability



Observe that $\frac{\partial f_1(\eta|\hat{\eta})}{\partial \hat{\eta}} > 0$ if $\hat{\eta} < \eta$, and $\frac{\partial f_1(\eta|\hat{\eta})}{\partial \hat{\eta}} < 0$ if $\hat{\eta} > \eta$. Hence, the second derivative is positive for low and negative for high values of $\hat{\eta}$. $E[\Pi_u(\tau, \eta)|\hat{\eta}]$ is first convex, then concave in $\hat{\eta}$. Note that $\lim_{\hat{\eta} \rightarrow -\infty} E[\Pi_u(\tau, \eta)|\hat{\eta}] = 0$: For a worker with very low expected ability, the probability of being more productive than the union wage is 0, independently of her training level. Furthermore, $\lim_{\hat{\eta} \rightarrow \infty} E[\Pi_u(\tau, \eta)|\hat{\eta}] = (1 - G(\Delta))\Delta$: For a worker with very high ability, the probability of being so productive that the union wage is not binding is 1, independently of her training level. Hence, when workers' expected ability is very low, profits without training also exceed profits with training by k . Similarly, when workers' expected ability is very high, profits without training also exceed profits with training by k . Consequently, there are two ability thresholds $\hat{\eta}_1$ and $\hat{\eta}_2$ such that expected profits with training exceed expected profits without training if $\hat{\eta}_1 < \hat{\eta} < \hat{\eta}_2$. See figure 3.4 for an illustration¹⁶.

We next show that unionised firms offer is less training than the socially optimal level. At the socially optimal level, $c'(\tilde{\tau}) = \frac{\partial E[y(\tilde{\tau}, \eta)|\hat{\eta}]}{\partial \tilde{\tau}}$, while the training level unionised firms offer satisfies $c'(\tilde{\tau}_u) = \frac{\partial E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}]}{\partial \tilde{\tau}_u}$. It is easy to see that $\frac{\partial E[y(\tilde{\tau}, \eta)|\hat{\eta}]}{\partial \tilde{\tau}} > \frac{\partial E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}]}{\partial \tilde{\tau}_u}$. Hence, unionised firms offer a lower training level than the socially optimal level.

Wage determination in the first period

It remains to analyse wage determination in the first period. Wages in the first period

¹⁶Note that if the fixed costs of training, k , are high, it may be optimal to train no worker.

are determined by the firm's zero profit condition. Since union agreements do not affect profits and wage determination in non-unionised firms, non-unionised firms offer the same first period wage as in the absence of union agreements, (3.2). Hence,

$$W_{nu} = W = \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}) + (1 - G(\Delta))\Delta.$$

The wage offer of unionised firms can be similarly derived as

$$W_u = \begin{cases} \min\{\bar{w}, \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}) + E[\Pi_u(0, \eta)|\hat{\eta}]\} \\ \quad \text{if } \hat{\eta} < \hat{\eta}_1 \text{ or } \hat{\eta} > \hat{\eta}_2 \text{ (no training).} \\ \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}) - k + E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}] - c(\tilde{\tau}_u) \\ \quad \text{if } \hat{\eta}_1 \leq \hat{\eta} \leq \hat{\eta}_2 \text{ (training).} \end{cases} \quad (3.8)$$

It is now apparent that firms bear the training cost. As the unionised firm only trains if the profit with training, $-k + E[\Pi_u(\tilde{\tau}_u, \eta)|\hat{\eta}] - c(\tilde{\tau}_u)$, exceeds the profit without training, $E[\Pi_u(0, \eta)|\hat{\eta}]$, the worker's training wage is higher than her first period wage would be without training.

C Worker sorting into the unionised sector

Non-unionised firms offer no training, while unionised firms offer τ_u^* . A worker chooses to work in the unionised sector if her utility from working in the unionised sector exceeds that from working in the non-unionised sector:

$$W_u(\tau_u^*, \hat{\eta}) + E[U_u(\tau_u^*, \eta)|\hat{\eta}] \geq W_{nu}(0, \hat{\eta}) + E[U_{nu}(0, \eta)|\hat{\eta}],$$

where $W_j(\tau_j, \hat{\eta})$, $j = u, nu$ denotes the worker's first period wage in a unionised or non-unionised firm, and $E[U_j(\tau_j^*, \eta)|\hat{\eta}]$, $j = u, nu$ denotes her second period utility when working in a unionised or non-unionised firm in the first period. We first derive the worker's second period utility when working in a non-unionised firm in the first period, $E[U_{nu}(0, \eta)|\hat{\eta}]$. If the worker leaves her employer, she is paid a wage equal to her productivity, $y(0, \eta)$. If she stays, her utility is equal to the wage the incumbent firm offers, $y(0, \eta) - \Delta$, plus the draw of non-pecuniary job characteristics, θ . The worker stays if $\theta > \Delta$. Hence, the worker's expected

utility in the second period equals

$$E[U_{nu}(0, \eta)|\hat{\eta}] = \int_{-\infty}^{\infty} y(0, \eta)dF_1(\eta|\hat{\eta}) + \int_{\Delta}^{\bar{\theta}} \theta - \Delta dG(\theta). \quad (3.9)$$

Using that the worker's first period wage equals $W_{nu}(0, \eta) = \int_{-\infty}^{\infty} y(0, \eta)dF_1(\eta|\hat{\eta}) + (1 - G(\Delta))\Delta$ (expression (3.2)), her utility from working in a non-unionised firm can be computed as

$$W_{nu}(0, \hat{\eta}) + E[U_{nu}(0, \eta)|\hat{\eta}] = 2 \int_{-\infty}^{\infty} y(0, \eta)dF_1(\eta|\hat{\eta}) + \int_{\Delta}^{\bar{\theta}} \theta dG(\theta). \quad (3.10)$$

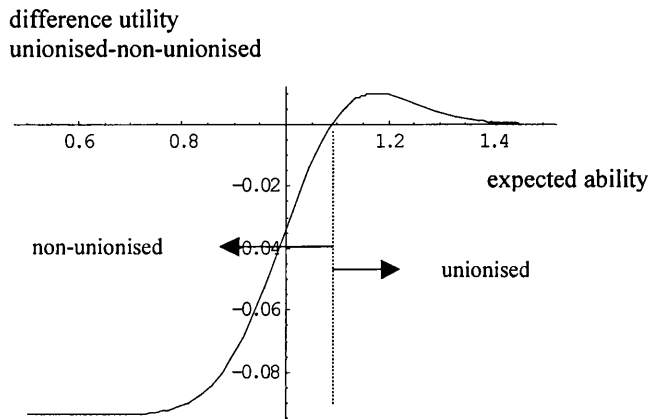
Next, we derive the worker's second period utility when working in a unionised firm in the first period, $E[U_u(\tau_u^*, \eta)|\hat{\eta}]$. Workers who turn out to be less productive than η_1 leave the unionised firm and are paid a wage equal to their productivity y . Workers whose ability is revealed to be between η_1 and η_2 get $\bar{w} + \theta$ if they stay and y if they leave. The probability that they stay is $1 - G(y - \bar{w})$. Finally, the utility of workers who turn out to be more able than η_2 is equal to y if they leave, and $y - \Delta + \theta$ if they stay. The probability of staying is $1 - G(\Delta)$. Hence, $E[U_u(\tau_u^*, \eta)|\hat{\eta}]$ can be computed as

$$\begin{aligned} E[U_u(\tau_u^*, \eta)|\hat{\eta}] &= \int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) + \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w}))\bar{w} dF_1(\eta|\hat{\eta}) + \int_{\eta_1}^{\eta_2} G(y - \bar{w})y dF_1(\eta|\hat{\eta}) + \int_{\eta_1}^{\eta_2} \int_{y-\bar{w}}^{\bar{\theta}} \theta dG(\theta) dF_1(\eta|\hat{\eta}) \\ &\quad + \int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) + (1 - F_1(\eta_2|\hat{\eta})) \int_{\Delta}^{\bar{\theta}} \theta - \Delta dG(\theta). \end{aligned} \quad (3.11)$$

Using expression (3.8) for the worker's wage in the first period, her utility from working in a unionised firm can be computed as

$$W_u(\tau_u^*, \hat{\eta}) + E[U_u(\tau_u^*, \eta)|\hat{\eta}] = \begin{cases} \begin{aligned} &2 \int_{-\infty}^{\infty} y(0, \eta)dF_1(\eta|\hat{\eta}) \\ &+ \int_{\eta_1}^{\eta_2} \int_{y(0, \eta) - \bar{w}}^{\bar{\theta}} \theta dG(\theta) dF_1(\eta|\hat{\eta}) \\ &+ (1 - F_1(\eta_2|\hat{\eta})) \int_{\Delta}^{\bar{\theta}} \theta dG(m) \end{aligned} & \text{if } \tau_u^* = 0, \\ \begin{aligned} &2 \int_{-\infty}^{\infty} y(0, \eta)dF_1(\eta|\hat{\eta}) - k - c() \\ &+ \int_{-\infty}^{\infty} y(\cdot, \eta)dF_1(\eta|\hat{\eta}) \\ &+ \int_{\eta_1}^{\eta_2} \int_{y(\cdot, \eta) - \bar{w}}^{\bar{\theta}} \theta dG(\theta) dF_1(\eta|\hat{\eta}) \\ &+ (1 - F_1(\eta_2|\hat{\eta})) \int_{\Delta}^{\bar{\theta}} \theta dG(m) \end{aligned} & \text{if } \tau_u^* > 0. \end{cases} \quad (3.12)$$

Figure 3.5: Difference between the utility from working in a unionised and non-unionised firm



The sorting of workers into unionised firms depends on the training level unionised firms offer as well as on their expected ability. First, consider the impact of training on worker sorting. Recall that unionised firms choose training such that the marginal cost of training is equal to the marginal profit of training: $c'(\tau_u) = \frac{\partial E[\Pi_u(\tau_u, \eta)|\hat{\eta}]}{\partial \tau_u}$. Also note that the worker's utility from working in a unionised firm can be alternatively be written as $W_u(\tau_u, \hat{\eta}) + E[U_u(\tau_u, \eta)|\hat{\eta}] = -k - c(\tau_u) + E[\Pi_u(\tau_u, \eta)|\hat{\eta}] + E[U_u(\tau_u, \eta)|\hat{\eta}]$. Hence, the training level that maximises the worker's utility,

Next, consider the impact of ability on worker sorting. Suppose unionised and non-unionised firms offer the same amount of training. Figure 3.5 plots the difference between the utility from working in a unionised and non-unionised firm as a function of workers' expected ability. First observe that from (3.12), the utility from working in a unionised firm converges to $2 \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta})$ as the worker's ability becomes very low. In contrast, from (3.10), the utility from working in a non-unionised firm converges to $2 \int_{-\infty}^{\infty} y(0, \eta) dF_1(\eta|\hat{\eta}) + \int_{\Delta}^{\bar{\theta}} \theta dG(\theta)$ as expected ability becomes low. Low ability workers are therefore better off in non-unionised firms. Second, observe that $\int_{y-\bar{w}}^{\bar{\theta}} \theta dG(\theta) > \int_{\Delta}^{\bar{\theta}} \theta dG(\theta)$: Workers whose ability turns out to be between η_1 and η_2 are better off due to unions. Workers with expected productivity around the union wage therefore prefer to work in the unionised sector. Finally, workers who turn out to be more able than η_2 are unaffected by the union

wage. Hence, as the worker's expected ability becomes very high, the difference between the utility from working in a unionised or non-unionised firm converges to 0. This implies that there exists an ability threshold such that workers with an expected ability above this threshold prefer to work in the unionised sector.

D Proof of proposition 3

Proposition 3 (i) In non-unionised firms, endogenous and exogenous movers are as productive as stayers and earn a higher wage than stayers. (ii) In unionised firms, endogenous and exogenous movers may be less productive and may earn a lower wage than stayers. (iii) The difference between the wage of endogenous movers and stayers is lower (in absolute terms) in unionised than in non-unionised firms.

(i) Stayers, exogenous and endogenous movers in non-unionised firms

Outside firms offer a wage equal to workers' productivity, i.e. $v = y(\tau, \eta)$, while incumbent firms offer a wage equal to workers' marginal productivity minus a constant, $w = y(\tau, \eta) - \Delta$. Hence, conditional on workers' productivity, wages of exogenous and endogenous movers exceed wages of stayers by Δ . Furthermore, recall that the probability of staying is $1 - G(v - w)$. As $v - w = \Delta$ for all workers, the staying probability does not depend on workers' productivity. Consequently, stayers, exogenous and endogenous movers have the same productivity on average.

(i) Stayers, exogenous and endogenous movers in unionised firms

We first compare the average productivity, and then the average wage, of stayers, exogenous and endogenous movers. For a worker with expected ability $\hat{\eta}$ the probability of staying with the unionised training firm is

$$Pr_u(\text{stay}|\hat{\eta}) = \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w}))dF_1(\eta|\hat{\eta}) + (1 - F_1(\eta_2|\hat{\eta}))(1 - G(\Delta)), \quad (3.13)$$

while the probability of moving can be computed as

$$Pr_u(\text{move}|\hat{\eta}) = F_1(\eta_1|\hat{\eta}) + \int_{\eta_1}^{\eta_2} G(y - \bar{w})dF_1(\eta|\hat{\eta}) + (1 - F_1(\eta_2|\hat{\eta}))G(\Delta).$$

The average productivity of stayers, conditional on expected ability, equals

$$E_u[y|\text{stay}, \hat{\eta}] = \frac{\int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) y dF_1(\eta|\hat{\eta}) + (1 - G(\Delta)) \int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta})}{Pr_u(\text{stay}|\hat{\eta})}, \quad (3.14)$$

while the average productivity of endogenous movers, conditional on expected ability, can be computed as

$$E_u[y|\text{move}, \hat{\eta}] = \frac{\int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) + \int_{\eta_1}^{\eta_2} G(y - \bar{w}) y dF_1(\eta|\hat{\eta}) + G(\Delta) \int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta})}{Pr_u(\text{move}|\hat{\eta})} \quad (3.15)$$

The average productivity of exogenous movers simply is

$$E[y|\text{exogenous}] = \int_{-\infty}^{\infty} y dF_1(\eta|\hat{\eta}).$$

Comparing the average productivity of stayers and endogenous movers, we get

$$\begin{aligned} & E_u[y|\text{stay}, \hat{\eta}] - E_u[y|\text{move}, \hat{\eta}] = \\ (a) \quad & \left\{ \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) y dF_1(\eta|\hat{\eta}) F_1(\eta_1|\hat{\eta}) - \right. \\ & \left. \int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) \right. \\ (b) \quad & \left. + (1 - G(\Delta)) \left(\int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} G(y - \bar{w}) dF_1(\eta|\hat{\eta}) - \right. \right. \\ & \left. \left. \int_{\eta_1}^{\eta_2} G(y - \bar{w}) y dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\ (c) \quad & \left. + (1 - G(\Delta)) \left(\int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) F_1(\eta_1|\hat{\eta}) - \int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\ (d) \quad & \left. - G(\Delta) \left(\int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) - \right. \right. \\ & \left. \left. \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) y dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\ (e) \quad & \left. - \left(\int_{\eta_1}^{\eta_2} G(y - \bar{w}) y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) - \right. \right. \\ & \left. \left. \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} G(y - \bar{w}) dF_1(\eta|\hat{\eta}) \right) \right\} / \\ & \{Pr_u(\text{stay})Pr_u(\text{move})\}. \end{aligned}$$

Terms (a) and (b) are clearly positive. Term (d) is negative, but adding also term (c) yields a positive expression, provided that the probability of staying exceeds the probability of

leaving. This clearly holds in the data. Term (e) is negative. To see this, observe that $\frac{\int_{\eta_1}^{\eta_2} G(y-\bar{w})y dF_1(\eta|\hat{\eta})}{\int_{\eta_1}^{\eta_2} G(y-\bar{w}) dF_1(\eta|\hat{\eta})} > \frac{\int_{\eta_1}^{\eta_2} (1-G(y-\bar{w}))y dF_1(\eta|\hat{\eta})}{\int_{\eta_1}^{\eta_2} 1-G(y-\bar{w}) dF_1(\eta|\hat{\eta})}$: For workers with ability between $\eta_1 < \eta < \eta_2$ it is the more able workers who are more likely to leave. Hence, conditional on workers' expected ability, it is ambiguous whether stayers are more or less able than movers. On the one hand, the least able workers are laid off ($\eta < \eta_1$). On the other hand, among workers who are paid the union wage ($\eta_1 < \eta < \eta_2$), more able workers are more likely to quit. By the same argument, it is ambiguous whether exogenous or endogenous movers are more productive.

Finally, note that as in non-unionised firms, exogenous and endogenous movers earn a higher wage than stayers conditional on productivity. Hence, a lower productivity of movers does not necessarily translate into a lower wage of movers. It is therefore ambiguous which group of workers earns the higher wage.

(iii) We have to show that the difference between the average wage of stayers and endogenous movers is greater than $-\Delta$. The average wage of stayers, conditional on expected ability, equals

$$E_u[w|\text{stay}, \hat{\eta}] = \frac{\int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w}))\bar{w} dF_1(\eta|\hat{\eta}) + (1 - G(\Delta)) \int_{\eta_2}^{\infty} y - \Delta dF_1(\eta|\hat{\eta})}{Pr_u(\text{stay}|\hat{\eta})},$$

while the average wage of movers is equal to the average productivity of movers, given by

(3.15). Taking the difference, we get

$$\begin{aligned}
& E_u[w|\text{stay}, \hat{\eta}] - E_u[v|\text{move}, \hat{\eta}] = \\
(a) \quad & \left\{ \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) \bar{w} dF_1(\eta|\hat{\eta}) F_1(\eta_1|\hat{\eta}) - \right. \\
& \left. \int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) \right. \\
(b) \quad & \left. + (1 - G(\Delta)) \left(\int_{\eta_2}^{\infty} y - \Delta dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} G(y - \bar{w}) dF_1(\eta|\hat{\eta}) - \right. \right. \\
& \left. \left. \int_{\eta_1}^{\eta_2} G(y - \bar{w}) y dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\
(c) \quad & \left. + (1 - G(\Delta)) \left(\int_{\eta_2}^{\infty} y - \Delta dF_1(\eta|\hat{\eta}) F_1(\eta_1|\hat{\eta}) - \int_{-\infty}^{\eta_1} y dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\
(d) \quad & \left. - G(\Delta) \left(\int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) - \right. \right. \\
& \left. \left. \int_{\eta_1}^{\eta_2} (1 - G(y - \bar{w})) \bar{w} dF_1(\eta|\hat{\eta}) (1 - F_1(\eta_2|\hat{\eta})) \right) \right. \\
(e) \quad & \left. (1 - G(\Delta)) G(\Delta) (1 - F_1(\eta_2|\hat{\eta})) \left(\int_{\eta_2}^{\infty} y - \Delta dF_1(\eta|\hat{\eta}) - \int_{\eta_2}^{\infty} y dF_1(\eta|\hat{\eta}) \right) \right. \\
(f) \quad & \left. - \int_{\eta_1}^{\eta_2} 1 - G(y - \bar{w}) dF_1(\eta|\hat{\eta}) \int_{\eta_1}^{\eta_2} G(y - \bar{w}) (y - \bar{w}) dF_1(\eta|\hat{\eta}) \right\} / \\
& \{Pr_u(\text{stay}) Pr_u(\text{move})\}.
\end{aligned}$$

Terms (a) and (b) are clearly positive. Term (d) is negative, but adding also term (c) yields a positive expression, provided that the probability of staying exceeds the probability of leaving. Term (e) and (f) are negative, but clearly smaller than $-\Delta$.

E Data

Worker data: Description, Sample Selection and definition of variables

The first data set we use is of administrative nature, drawn from the German Social Security Records, the so-called IAB-Beschäftigtenstichprobe. It constitutes a one percent sample of the German labour force subject to social security contributions for the period between 1975 to 1995. The data contains an unusual set of background information for the individual, including age and education. It allows to construct the complete work history, including time spent in registered unemployment, from labour market entry onwards up to 20 years in

the labour force. Furthermore, due to the administrative nature of this data set, wages and employment spells are very accurate, and, other than in many surveys, wages are uniquely assigned to particular jobs.

Firms report to the Federal Employment Office any beginning and end of an employment relationship when these occur, as well as information on ongoing relationships at December 31st every year. The wage reported in our data is the average daily wage computed for each spell. We deflate wages using the Consumer Price Index, with 1995 as the basis year. Actual experience is measured as weeks (divided by 52) spent in full-time employment from the 1st of January following the year the apprenticeship ended. Part time employment, time as apprentice, time spent unemployed as well as time spent out of the labour force is not counted. We assume that the wage data refers to experience accumulated at the start of each employment spell.

A disadvantage of the data set is that it does not cover the entire German labour force, but only employees within the social security net. This excludes, among others, the self-employed and civil servants. As many administrative data sets, our data is right-censored at the highest level of earnings that are subject to social security contributions. In our sample top-coding is not a serious problem: Less than 0.5 % of all wage observations are top-coded.

We select from this data base a sample of male individuals who all went through apprenticeship training. We restrict analysis to former West Germany only. We define an apprentice as a worker who is reported as an trainee in one company for at least 700 days, no matter whether the education variable classifies him as a worker with or without a finished apprenticeship. In addition, we consider a worker with a work placement as an apprentice if the work placement lasted at least 1 year (position variable) and he was employed as a worker with a completed apprenticeship in at least half of his jobs (education variable).

Sample selection An apprentice is included in the sample if

- he completed the apprenticeship after 1979;
- he has never worked in East Germany;
- he completed only one apprenticeship;

- if he was not older than 15 in 1975 and at most 19 at labour market entry and had completed no A-levels (Abitur) at labour market entry;
- if he was not older than 15 in 1975 and at most 21 at labour market entry and had completed A-levels (Abitur) at labour market entry.

We end up with a sample of 24194 individuals.

Our data set allows us to precisely determine whether a worker has left the apprenticeship firm at the end of the training period. However, for apprentices who remain with the same firm after training a separation between training spell and employment spell in the year when training is concluded is not possible. When we compare entry wages of stayers and movers, we therefore assign to both stayers and movers the first wage spell in the year following the year the training period ended.

We define two types of exogenous movers. First, we follow Acemoglu and Pischke (1998) and use draft into the military service or social service as a measure for exogenous separation from the training firm¹⁷. We refer to this variable as service. Second, our data set contains also information on the size of the firm, as of June of the respective calendar year¹⁸. From this information we construct a measure for firm closure by computing the year in which establishment size drops to zero.

Military and civil service last between 10 and 30 months, depending on the type of service, and the year the service started. For most of the workers in our data set, service lasted between 10 and 18 months. Our data does not contain direct information on whether workers are doing military or civil service. We approximate service by the time workers spend neither in registered employment nor in registered unemployment.¹⁹

We use three alternative definitions. According to our first definition workers are classified as being in military or civil service if they are registered neither as employed nor as

¹⁷See Acemoglu and Pischke (1998) for details on compulsory service schemes in Germany.

¹⁸This information has been created by aggregating up the number of individuals for each firm identifier by year from the population of individuals in the 100% sample.

¹⁹Workers who have finished apprenticeship training are entitled to unemployment benefits if they register as unemployed. Hence, it is likely that unemployed workers are also registered as unemployed.

Table 3.6: Military and Social Service: 3 Definitions

	Definition 1	Definition 2	Definition 3
total workers	24194	24194	24194
military service	6611	6882	8074
no military service	13516	13516	13430
problematic cases	4067	3796	2690
Σ	24194	24194	24194
leaves due to service	2459	2510	2941
dos not leave due to service	4152	4372	5113
Σ	6611	6882	8074
problematic cases	1998	1923	1482
goes back to appr. firm	778	805	895
goes to new firm	1681	1705	2046
Σ	2459	2510	2941

unemployed for at least 10 consecutive months and not longer than 18 months. Some workers in military service are still registered with their previous employer. These workers typically return to the previous employer after the military service. For these workers, firms report the working relationship as interrupted. We also classify these workers as in military service if the time between the interruption spell and the beginning of the new job is at least 10 months and at most 18 months. Workers are classified as not doing their military or civil service if they have no non-employment gap longer than 10 months and no interruption spell longer than 10 months. If workers have more than one non-employment gap or interruption spell longer than 10 months, they are classified as "problematic". We also consider a worker as problematic if both the interruption spell and the non-employment gap identifies the worker as in military service.

Our second definition follows definition 1, but workers who are classified as in service according to both the interruption spell and the non-employment gap are classified as in service, and not as problematic.

Definition 3 follows definition 2, but extends the duration of the non-employment gap

and the interruption spell to 30 months.

Table 3.6 shows the proportion of workers in service for the three definitions. According to definition 1, 6611 out of 24194 workers do their military or civil service during the observation period. Of those, 2449 leave the apprenticeship firm due to service. Of those, 778 return to the apprenticeship firm. 4067 workers are classified as problematic. However, for our analysis it is only important whether the worker is classified as problematic in the year he finished his apprenticeship. This is the case for 1998 workers. A similar decomposition for definition 2 and 3.

To construct separations due to plant closure, we use information on the size of the firm as of June 1st each year. We classify a worker as an exogenous mover due to plant closure if the size of the training firm drops to zero in the year the worker finishes the apprenticeship training period. Furthermore, if the worker completed the training period after June 1st, he is classified as an exogenous mover if the size of the training firm drops to zero in the next year.

Firm data

Our firm data is based on a yearly panel of establishments, the so-called IAB-Betriebspanel. The base population are all firms with at least one employee who pays social security contributions. In 1993 the panel started with about 4,000 establishments; in 2000 around 13,000 firms participated in the survey (see Kölling (2000)). Large firms are over-sampled. We restrict our analysis to West German firms in the private sector only.

Information on firms is collected through person-to-person interviews with the firm management. In addition to the union status, the data contains a large array of background characteristics, including firm size, industry, investment, revenue, etc. A firm is considered as "unionised" if it belongs to an employer association and is bound to the industry agreement that is relevant for their particular firm (Branchentarifvertrag), or if it engages in bilateral negotiations with the union (Haustarifvertrag)²⁰.

²⁰See appendix E for a brief description of the German collective bargaining system. In our sample, the proportion of workers who are employed in firms that engage in firm-level bargaining is 9 % (IAB-Betriebspanel, 1996-1999, worker weights).

We are able to match for each firm and year average worker characteristics, which allows us to compute the fraction of apprentices in the firm. Average worker characteristics are computed from the social security records for each firm in our establishment sample as of June 1st each year. This data is available from 1996-1999. We therefore use only the data from 1996-1999 for the empirical analysis. We end up with an unbalanced panel of 4717 firms and 13453 observations.

Table 3.7 lists and defines the variables of the firm data used in the empirical analysis.

Table 3.7: Variable definitions: Firm data

variable	definition
union status	1 if firm is bound to industry or firm level agreement, 0 otherwise
firm size	total number of employees in the firm
industry	10 industry dummies: Energy/mining/water industry; chemical industry; metal industry/machines; electro-technical industry/automobiles/optical industry; wood/printing/paper; construction/capentry; retail/wholesale; traffic/news; credit/insurance; other services agriculture, charities/private households and public sector dropped
revenue/worker	total turnover in the firm in the previous year divided by number of employees
investment/worker	total sum of investments in the previous year divided by number of employees
age of firm	distinguishes between 5 years and younger, 6-15 years, 16-30 years, older than 30 years
evaluation of profit	firm's current evaluation of profit; from 1 (very good) to 5 (very bad)
proportion apprentices	number of apprentices divided by number of employees
firm trains	1 if firm employs at least one apprentice, 0 otherwise

Matched data

It is in principle possible to match information on workers to the firm data. As our worker data, the information on workers comes from social security records. In addition to average worker characteristics for each firm for the years 1996-1999, we have access to the following 'matched' data set. For the comparison of the mover-stayer wage differential of workers trained in unionised and non-unionised firms we select all workers who complete an apprenticeship in one of the firms in the establishment panel. A worker is included in the sample if the apprenticeship ended between 1995 and 1997 and was done in the private sector in a West German firm. For each worker we observe the first post-apprenticeship wage spell, whether he stayed with or left the training firm, as well as the same background characteristics as in the worker data. We match to this data the union status of the training firm from the establishment panel.

Chapter 4

Job Mobility in the United States and Germany

The American labour market is generally considered as one of the most flexible labour markets in the advanced world. The German labour market, in contrast, is a regulated labour market with severe firing restrictions¹ and centralised wage-setting institutions. It is also well known that German workers switch jobs less often than American workers. For instance, Topel and Ward (1992) find that after ten years American workers are on their 7th job on average, while German workers have worked for only 3 jobs (Dustmann and Meghir (2003)).

This chapter provides new evidence on the job search process of young men in both countries. Based on a search model with endogenous job-to-job and job-to-unemployment transitions, we provide a simple method of decomposing the variance of the unexplained wage component into the variance of ability, match quality, and transitory component by experience.

While there are several empirical studies on job mobility for the US², studies for

¹According to ILO rankings, on a scale from 0 (firing costs non-existent) to 3 (firing costs severe) the US ranks 0.4 and Germany 2.7 (OECD (1999)). Only Italy, Spain and Portugal are ranked higher than Germany.

²Studies on job mobility in the US include Topel and Ward (1992), Farber (1994), and Light and McGarry (1998).

Germany are more rare³. The empirical analysis for Germany is based on a new and unique administrative data set, the so-called IAB-Beschäftigtenstichprobe. The data for the US comes from the 1979 National Longitudinal Survey of Youth (NLSY79). Both data sets are well suited to analyse the mobility behaviour of young workers. Most importantly, both data sets allow to construct an accurate and complete work history of workers from labour market entry onwards.

The search model borrows from Jovanovic (1984) and Mortensen (1988). The two crucial features of the model are on-the-job search and learning about match quality. On-the-job search leads to endogenous job-to-job transitions and allows workers to allocate better matches over time. Learning about the quality of the match leads to endogenous job-to-unemployment transitions. A match that used to be preferable to unemployment may not be preferable anymore after disappointing news about the match arrived. A distinction between job-to-job and job-to-unemployment mobility is important. When moving from job-to-job workers gain search capital, whereas they lose search capital when becoming unemployed. We extend Mortensen's (1988) and Jovanovic's (1984) model by allowing the worker's productivity to depend not only on the quality of the match, but also on ability. We incorporate ability in the simplest possible way, and assume that ability does not affect the worker's decision to switch jobs. We do this in order to focus on how match quality and learning about match quality affects mobility.

Our main findings can be summarised as follows. Mobility is substantially lower in Germany than in the US for all education groups. German workers are particularly less likely to become unemployed. Yet, there are important similarities between US and German workers. Match-specific productivity plays an important role in both countries for all education groups. As predicted by the search model, the autocovariance of log-wage residuals is highest for stayers, second highest for job-to-job movers and

³The focus of Dustmann and Meghir (2003) is on estimating the returns to experience and tenure. Special emphasis is given to how job mobility affects the estimation of returns to experience and tenure. Winkelmann (1996a, 1996b) compare mobility rates and overall wage growth of apprentices, university graduates, and workers without post-secondary education.

lowest for job-to-unemployment movers. Furthermore, the proportion of the variance that can be attributed to match declines with experience, while the proportion that can be attributed to ability increases with experience in both countries for all education groups. There is also some evidence for learning about match quality in both countries, as the variance of the within-job wage growth residual declines with tenure. Heterogeneity in the rate of firm-specific human capital accumulation, in contrast, is not an important feature of the data in either country, as the autocovariance of the within-job wage growth residual is close to zero and uniformly negative at lag greater than 1.

American and German workers mostly differ with respect to the variance of ability and transitory wage component. As expected, the variance of the transitory wage component is substantially lower in the German administrative data than in the US survey data due to measurement error. The variance of ability is considerably higher in US for all education groups. Furthermore, learning about ability plays an important role in the US, but only a modest role in Germany. The proportion of the variance attributable to match quality is considerably higher in Germany than in the US. This may indicate that search frictions play a more important role in Germany.

There are also interesting differences between unskilled workers and apprentices in Germany. First, the variance of the quality of the match is almost three times as high for unskilled workers than for apprentices at low levels of experience. This difference all but disappears with experience. This is consistent with the idea that job search plays a more important role for unskilled workers than for apprentices. Second, the variance of the within-job wage growth residual is considerably higher for unskilled workers than for apprentices at low tenure levels, particularly so at the first job. This difference all but disappears with tenure. One reason for the lower mobility of apprentices than for unskilled workers may thus be a higher precision of the quality of the match.

The structure of this chapter is as follows. Section 4.1 describes the search model and analyses the worker's mobility decision. Section 4.2 derives the empirical implications of the search model, and proposes a simple method of decomposing the variance of the log-wage residual into the variance of ability, match quality and a transitory

component. Section ?? discusses alternative explanations for the lower mobility in Germany within the search model. Section 4.3 describes the data. Section 4.4 compares mobility as well as the wage gains from mobility in the US and Germany. The variance of ability, match quality and the transitory component is estimated. Section ?? concludes.

4.1 A model of job-to-job and job-to-unemployment mobility

In this section we set up a search model with endogenous job-to-job and job-to-unemployment transitions. We borrow from Jovanovic (1984) and Mortensen (1988). The two crucial features of the model are on-the-job search and learning about match quality. On-the-job search leads to endogenous job-to-job transitions and allows workers to allocate better matches with time in the labour market. Learning about match quality leads to endogenous job-to-unemployment transitions. We extend Jovanovic's (1984) and Mortensen's (1988) model by allowing the worker's productivity to depend not only on the quality of the match, but also on ability. We incorporate ability in the simplest possible way, and assume that ability does not affect the worker's decision to switch jobs. We do this in order to focus on how match quality and learning about match quality affects mobility and wage dynamics. Next, we describe the model in more detail (section 4.1.1). We then characterise the worker's mobility decision (4.1.2).

4.1.1 Description of environment

The model is set in continuous time. Workers live forever. They are risk-neutral, and maximise the expected present value of life-time income. This implies that utility of a worker is equal to his wage. There is no borrowing or saving.

Productivity The worker's productivity depends on his ability, denoted by f , the quality of the match, denoted by μ , as well as on firm-specific human capital accumu-

lation. In the empirical analysis, we also allow for general human capital accumulation as well as aggregate productivity shocks. We specify worker i 's productivity in firm j as follows.

$$y_{ij} = e^{f_i} e^{h(n) + \mu_{ij}}.$$

Here, $h(n)$ represents firm-specific human capital accumulation when tenure is equal to n (see below for a precise definition of tenure). We assume that h is concave, i.e. workers accumulate more firm-specific human capital early on-the-job than later. We think of firm-specific human capital accumulation as learning-by-doing, and not as an investment decision. Firm-specific human capital accumulation is independent of both ability and the quality of the match⁴. The distribution of the quality of the match is normal with mean $\bar{\mu}$ and precision $1/\sigma_\mu^2 = \tau_\mu$. It is independent of tenure and ability. Let \bar{f} and $\tau_f = 1/\sigma_f^2$ denote the mean and precision of the worker's ability distribution.

Unemployment Worker enters the labour market as unemployed. They meet a firm at a given Poisson frequency equal to λ_0 . Upon meeting they decide whether to accept or reject the job offer. There is no recall. The worker's flow utility from unemployment is $e^{f_i} e^b$. The worker must forego e^{f_i+b} when employed. It therefore represents the income-equivalent of the disutility from work. This particular specification about the disutility from work rules out that ability has an effect on the probability of accepting a job when unemployed or on the probability of becoming unemployed - see section 4.1.1⁵. This is the first restriction we impose in order to decompose the variance of the log-wage residual into the variance of ability, match quality, and a transitory component.

R.1 Ability does not affect the worker's mobility decision.

Screening When firms and workers meet, firms screen the worker and receive a

⁴Jovanovic (1979b) sets up a model in which workers optimally invest more in firm-specific human capital if the quality of the match is higher. Neal (1998) develops a model in which more able workers have a comparative advantage at investing in specific as opposed to general training.

⁵Postel-Vinay and Robin (2002) impose a similar restriction.

noisy signal about the *quality of the match*. The signal m equals

$$m = \mu + \epsilon_m,$$

with $\epsilon_m \sim N(0, \sigma_m^2)$. Let $\tau_m = \frac{1}{\sigma_m^2}$ denote the precision of the screen. Firms and workers use the signal to update their belief about the match quality. The smaller σ_m^2 , the better the screen. If $\tau_m \rightarrow 0$, then nothing can be learned about the match before output takes place. The match is a pure experience good. If, in contrast, $\tau_m \rightarrow \infty$, a match is a pure inspection good and its quality can be fully ascertained by inspection. Bayesian updating implies that the posterior belief about the quality of the match is normally distributed with mean (DeGroot (1970))

$$E[\mu|m] = \frac{\bar{\mu}\tau_\mu + m\tau_m}{\tau_\mu + \tau_m} := \mu_0, \quad (4.1)$$

and variance

$$S_0 = \frac{1}{\tau_\mu + \tau_m}. \quad (4.2)$$

The updated belief about the quality of the match is a weighted average of the prior mean, $\bar{\mu}$, and the signal, m . The more weight is given to $\bar{\mu}$ the higher τ_μ , i.e. the more precise the prior belief about productivity. The more weight is given to m the higher τ_m , i.e. the more precise the signal. If the match is a pure experience good, i.e. $\tau_m \rightarrow 0$, then $E[\mu|m] = \bar{\mu}$ and $S_0 = \sigma_\mu^2$. If, on the other hand, the match is a pure inspection good, i.e. $\tau_m \rightarrow \infty$, then $E[\mu|m] = m$ and $S_0 = 0$.

We can also compute the distribution from which the expected match quality after screening, μ_0 , is drawn. In appendix D we show that it follows a normal distribution with mean $\bar{\mu}$ and variance

$$V(\mu_0) = \frac{\tau_m}{\tau_\mu(\tau_m + \tau_\mu)}. \quad (4.3)$$

In the case of a pure inspection good, i.e. $\tau_m \rightarrow \infty$, $V(\mu_0)$ converges to the actual variance of the quality of the match, $1/\tau_\mu$. In the opposite case of a pure experience good, i.e. $\tau_m \rightarrow 0$, all jobs look the same ex ante and $V(\mu_0) = 0$.

As the focus of the search model is how match quality and learning about match quality affect mobility, we assume that firms and workers are perfectly informed about

ability. This assumption can be relaxed by the assumption that firms and workers symmetrically learn about ability. We comment on how symmetric employer learning affects the empirical implications. In particular, we discuss how the variance of ability, match quality and a transitory component can be recovered from the data under symmetric learning about ability.

Evolution of beliefs about match quality New information about the quality of the match arrives on-the-job when the worker 'experiences' the job. This may lead to job-to-unemployment transitions. A match that used to be preferable to unemployment may not be preferable anymore after disappointing news about the match arrived. Jobs are also destroyed for exogenous reasons, for instance because of plant closure. Let δ denote Poisson frequency of the exogenous job destruction rate. We assume that a new signal about the quality of the match arrives at a given Poisson frequency equal to θ . The signal about the quality of the match, x , is noisy, and equals

$$x = \mu + \epsilon_x,$$

with $x \sim N(0, \sigma_x^2)$. Let $\tau_x = \frac{1}{\sigma_x^2}$ denote the precision of the signal. For a match with screen outcome m , job tenure n , and average signal \bar{x} , the posterior distribution of the quality of the match is normal with mean

$$E[\mu|\bar{x}, m, n] = S_n(\tau_\mu \bar{\mu} + \tau_m m + \tau_x n \bar{x}) := \mu_n, \quad (4.4)$$

and variance

$$S_n = \frac{1}{\tau_\mu + \tau_m + n\tau_x}. \quad (4.5)$$

As tenure increases, the posterior variance decreases. The belief about the quality of the match becomes increasingly precise.

Search on-the-job Workers continue to search for a better job when employed. Outside offers arrive at a Poisson frequency equal to λ_1 . When an employed worker meets a firm, the firm screens the worker. Firms use the screen to update the belief about the quality of the match and make a wage offer based on the screening outcome. The

worker decides whether to accept the outside wage offer and move from job-to-job or whether to reject the outside offer and remain with his current employer.

Wage determination The worker's wage at tenure n is equal to the worker's expected marginal product conditional on the screen m and the average signals received on-the-job, \bar{x} :

$$w_{ij} = E[e^{f_i} e^{h(n)+\mu_{ij}} | \bar{x}, m, n].$$

Each time new information about the worker's productivity arrives, the firms makes a new wage offer to the worker. The wage offer reflects both the new information about the quality of the match and the increase in firm-specific human capital, $h(n) - h(n - 1)$. Here, n correctly represents the current number of signal realisations. We refer to n as tenure. A justification for this terminology is that the number of signal realisations received on-the-job can be easily transformed into the time spent with the same employer (see also Mortensen (1988)). Jovanovic (1979a) shows that this an equilibrium wage contract if firms bid for workers by offering life-time contracts which they honour ex post⁶.

4.1.2 The worker's mobility decision

Mortensen (1988) and Jovanovic (1984) characterise the worker's optimal mobility decision under similar assumptions. The most important difference is that in Mortensen (1988) and Jovanovic (1984) the worker's productivity only depends on the quality of the match, but not on his ability. We have incorporated ability into the search model

⁶A problem with this type of contract is that it is not unique. Moreover, it is not self-enforcing. As the worker's productivity in the current firm differs from his market productivity, the worker may stay even if he is not paid his full marginal product. Our approach to this problem is as follows. The worker's mobility behaviour is essentially unchanged if separation is efficient; that is, a job is terminated only if the total value of the job, i.e. the sum of the worker' and the firm's value of the job, is lower than the sum of the worker's and firm's outside option. Our results therefore also hold if workers and firms share the total surplus of the match, and wages are determined by Nash bargaining. Moreover, our empirical results provide strong evidence for the importance of match- or firm-specific productivity. This implies that there is scope for job search in the economy.

in the simplest way. Our assumption on the flow utility from unemployment implies that ability does not affect the worker's productivity decision. To see this, let U , W and W' denote the value of unemployment, the value of the current job and the value of the outside job, respectively, for a worker with average ability \bar{f} . Then the value of unemployment, the value of the current job and the value of the outside job for a worker with ability f_i is equal to $e^{f_i - \bar{f}}U$, $e^{f_i - \bar{f}}W$, and $e^{f_i - \bar{f}}W'$, respectively. Since the worker chooses the state with the maximum value, ability has no impact on the worker's mobility decision.

Mortensen (1988) and Jovanovic (1984) show that the worker's mobility decision only depends on tenure and the wage he is offered by her current firm and the outside firm. Since in our case the worker's wage does not only depend on the quality of the match but also on his ability, we need to make a small adjustment. Recall that the worker is paid his expected marginal product, i.e. $w = e^f E[e^{h(n)+\mu}|X, m, n,]$. Define $\ln \tilde{w}$ as the log-wage that corrects for the worker's ability, i.e. $\ln \tilde{w} = \ln(\frac{w}{e^f}) = \ln E[e^{h(n)+\mu}|X, m, n,]$. The worker's mobility decision is then fully characterised by his corrected log-wage $\ln \tilde{w}$ and tenure n .

Unemployment-to-job transition Consider first the decision of an unemployed worker to accept a job offer. It satisfies the reservation property. Let $\Psi(0)$ denote the reservation wage that makes the worker indifferent between accepting the job offer and remaining unemployed. Suppose the worker's corrected log-wage at the employer is $\ln \tilde{w}$. The worker's decision rule is

- accept if $\ln \tilde{w} \geq \Psi(0)$,
- reject if $\ln \tilde{w} < \Psi(0)$.

Since the worker accumulates firm-specific human capital and is only imperfectly informed about the quality of the match, he is willing to accept a job offer which pays a lower wage than the flow utility from unemployment, i.e. $\Psi(0) < b$.

Job-to-unemployment transition Next, consider the decision of a worker with tenure n to become unemployed. It again satisfies the reservation property. Let $\Psi(n)$

denote the reservation wage that makes the worker indifferent between staying with the current employer and becoming unemployed. Suppose the worker's corrected log-wage at the current firm is $\ln \tilde{w}$. The worker's decision rule is

- become unemployed if $\ln \tilde{w} < \Psi(n)$,
- stay if $\ln \tilde{w} \geq \Psi(n)$.

Mortensen (1988) and Jovanovic (1984) show that the reservation match quality is increasing in tenure, i.e. $\Psi(n) \geq \Psi(n-1)$ for all n . This is so for two reasons. The first reason is firm-specific human capital accumulation. Since workers accumulate more firm-specific capital early on-the-job, jobs with lower tenure provide the opportunity of a higher wage growth. The option value of the job therefore decreases as tenure increases. The worker therefore requires a higher wage in order to stay as tenure increases. Since the worker always has the option of becoming unemployed, he is insured against downside risk. Hence, uncertainty about the quality of the match offers the possibility of future wage growth. As the estimate of the quality of the match becomes increasingly precise with tenure, the potential for future wage growth decreases with tenure. The reservation wage therefore increases with tenure.

Job-to-job transition Finally, consider the worker's decision to move from job-to-job after n periods of tenure. Suppose the worker's corrected log-wage is $\ln \tilde{w}$ at his current firm and $\ln \tilde{w}'$ at the outside firm. The worker's decision rule follows a reservation policy. Let $\rho(\ln \tilde{w}, n)$ denote the reservation offer that makes a worker indifferent between staying with the current employer and switching to the new employer. The worker's decision rule is

- move from job-to-job if $\ln \tilde{w}' > \rho(\ln \tilde{w}, n)$,
- stay if $\ln \tilde{w}' \leq \rho(\ln \tilde{w}, n)$.

Jovanovic (1984) and Mortensen (1988) show that the reservation match quality is increasing in the wage, i.e. $\frac{\partial \rho(\ln \tilde{w}, n)}{\partial \ln \tilde{w}} \geq 0$. Intuitively, well paid workers require a higher wage than badly paid workers in order to switch jobs. Jovanovic (1984) and Mortensen (1988) also show that the reservation offer decreases with tenure conditional on the wage, i.e. $\rho(\ln \tilde{w}, n) \leq \rho(\ln \tilde{w}, n-1)$ for all n . The reason is the same as why the

reservation wage $\Psi(n)$ increases with tenure. First, the new job promises a higher wage growth due to firm-specific human capital accumulation. Second, the uncertainty of the match quality and thus the potential for future wage growth is higher at the new than at the current job. This means that the worker is willing to accept a wage cut in exchange for a higher wage growth when switching jobs.

It is important to emphasise that the reservation wage and offer -and thus the worker's mobility decision- does not depend on experience. Thus, the search problem of an unemployed with ten years of experience is the same as that of a worker who just entered the labour market. The search process starts afresh each time the worker becomes unemployed. This is the second restriction we impose in order to decompose the variance of the log-wage residual into the variance of ability, match quality and a transitory component.

R.2 The worker's search problem does not depend on experience.

Unemployed workers have to search from scratch.

There are several reasons why this restriction may be violated. First, workers may be entitled to unemployment benefits only after they accumulated experience. Moreover, the unemployment benefit may depend on the worker's previous wage. Second, workers may accumulate more general human capital when young so that the option value of a new job decreases with experience. More experienced workers will thus be "choosier" and reject job offers when unemployed which labour market entrants would have accepted. Finally, Neal (1999) proposes a search model in which workers first search for a good occupation-specific match and then for a good firm-specific match. As the occupation-specific search capital is portable to other firms within the same occupation, his model implies that workers do not lose the entire search capital when unemployed.

4.2 Empirical implications

In this section we derive some empirical implications of the search model. We begin with the gains from mobility by analysing the wage gains of stayers, job-to-job and job-to-unemployment movers (section 4.2.1). We then derive the restrictions the search model imposes on the covariance structure of log-wage residuals, and propose a simple method of decomposing the variance of the log-wage residual into the variance of ability, the variance of the transitory wage component and the variance of the quality of the match (section 4.2.2). This provides further evidence for the importance of match-specific productivity and job search.

4.2.1 Wage gains by mobility groups

Suppose that for each worker we observe a sequence of wages. We index the sequence by time t . It is possible that we do not observe all wage changes a worker experiences. To simplify the notation, we assume that stayers have experienced exactly one wage adjustment between the two dates the wage data is available (i.e. between t and $t + 1$). Furthermore, we assume that for job-to-job and job-to-unemployment movers the wage at the old job did not change before they switched jobs.

Recall that wages are equal to expected productivity, i.e. $w = E[e^f e^{h(n)+\mu} | \bar{x}, m, n]$. Also recall that $E[\mu | \bar{x}, m, n]$ is normally distributed with mean μ_n and variance S_n . Hence, $w = e^f e^{h(n)+\mu_n+0.5S_n}$. Although the actual quality of the match does not change on-the-job, its perception does. We therefore index the quality of the match by tenure n . Adding an i.i.d. error term ε , the log-wage of worker i in firm j at time t in can be written as

$$\ln w_{ijt} = f_i + h(n) + 0.5S_n + \mu_{ijn} + \varepsilon_{ijt}. \quad (4.6)$$

In principle, ε may reflect both measurement error and transitory wage shocks. We assume that ε does not affect the worker's mobility decision. This is the third restriction we impose in order to decompose the variance of the log-wage residual into the variance

of ability (f), match quality (μ) and the transitory component (ε).

R.3 ε does not affect the worker's mobility decision.

Specification (4.6) assumes that firms are perfectly informed about workers' ability. Later we also comment on the case in which firms symmetrically learn about ability.

Wages grow for two reasons. First, workers accumulate firm-specific human capital. Second, workers become better matched as they spend more time in the labour market. We refer to this source of wage growth as wage growth due to search. Next, we derive the wage gains of stayers, job-to-job and job-to-unemployment movers.

We have ignored wage growth due to general human capital accumulation in order to focus on how match quality and learning about match quality affects mobility and wage dynamics. In the empirical analysis, we allow for general human capital accumulation.

Wage gains of stayers Taking the first difference of (4.6), the average wage gain conditional on staying (s) equals

$$E[\Delta \ln w_{ijt}|s] = \Delta h(n) + 0.5\Delta S_n + E[\Delta \mu_{ijn}|s].$$

Bayesian updating implies that the quality of the match follows a martingale without drift, i.e.

$$\mu_n = \mu_{n-1} + \eta_n,$$

with $E[\eta_n] = 0$. In words, the best guess for the worker's productivity tomorrow is his productivity today. We can therefore rewrite the average wage gain of stayers as

$$E[\Delta \ln w_{ijt}|s] = \Delta h(n) + 0.5\Delta S_n + E[\eta_{ijn}|s].$$

Since jobs that turned out to be a disappointment are terminated, $E[\eta_{ijn}|s] \geq 0$ although $E[\eta_{ijt}] = 0$. Consequently, the wage growth of stayers identifies the sum of the wage growth due to firm-specific human capital, $\Delta h(n)$, and the wage growth due to learning about match quality, $0.5\Delta S_n + E[\eta_{ijn}|s]$. Hence, under learning about match quality part of the wage growth due to search shows up as within-job wage growth.

Wage gains of job-to-job movers Next, consider the wage growth of workers who move from job-to-job (from firm j to firm j'). The wage gain of job-to-job movers (j) equals

$$E[\Delta \ln w_{ijt}|j] = -h(n-1) + 0.5\Delta S_n + E[\mu_{ij'0} - \mu_{ijn-1}|j].$$

Here, $h(n-1)$ represents the loss in firm-specific human capital, while $0.5\Delta S_n + E[\mu_{ij'n} - \mu_{ijn-1}|j]$ captures the gain in search capital when switching jobs. Note that wage growth due to search shows up both as within-job and job-to-job wage growth. Workers realise gains from job search not only when switching jobs, but also when staying with the same employer (see also Jovanovic (1984)).

Wage growth of job-to-unemployment movers Finally, the wage change of job-to-unemployment movers (u) equals

$$E[\Delta \ln w_{ijt}|u] = -h(n-1) + 0.5\Delta S_n + E[\mu_{ij'0} - \mu_{ijn-1}|u].$$

Job-to-unemployment movers lose their search capital in addition to their firm-specific human capital.

4.2.2 The covariance structure of log-wage residuals

In this section we derive the restrictions the search model imposes on the covariance structure of log-wage residuals. We propose a simple method of decomposing the variance of the log-wage residual into the variance of ability (σ_f^2), the variance of measurement error (σ_ϵ^2) and the variance of the quality of the match. This provides further evidence on the importance of match-specific productivity and gains from job search. We first analyse the covariance structure of log-wage residuals separately for workers who stay, move from job-to-job and from job-to-unemployment and show how the variance of ability can be identified. We then analyse how the distribution of the quality of the match varies with experience. Finally, we analyse the covariance structure of the within-job wage growth residual and show how the variance of the transitory wage component can be identified.

The covariance structure of log-wage residuals by mobility groups: Identifying the variance of ability

From (4.6), the log-wage of worker i in firm j at time t equals $\ln w_{ijt} = f_i + h(n) + 0.5S_n + \mu_{ijn} + \varepsilon_{ijt}$. Define u_{ijt} as the residual of the log-wage regression, i.e.

$$u_{ijt} = f_i + \mu_{ijn} + \varepsilon_{ijt}.$$

Consider the autocovariance of the residual between t and $t - 1$. The search model predicts that the autocovariance depends on whether the worker stayed with his employer, moved from job-to-job or from job-to-unemployment between t and $t - 1$. The autocovariance at lag 1 for stayers (s), job-to-job movers (j), and job-to-unemployment movers (u) can be computed as⁷

$$\begin{aligned} Cov(u_{ijt}, u_{ijt-1}|s) &= \sigma_f^2 + Cov(\mu_{ijn}, \mu_{ijn-1}|s), \\ Cov(u_{ijt}, u_{ijt-1}|j) &= \sigma_f^2 + Cov(\mu_{ij'0}, \mu_{ijn-1}|j), \text{ and} \\ Cov(u_{ijt}, u_{ijt-1}|u) &= \sigma_f^2. \end{aligned}$$

First, consider workers who stay with their employer. As $\mu_n = \mu_{n-1} + \eta_n$, the covariance of the quality of the match at tenure n and $n - 1$, conditional on staying, equals

$$\begin{aligned} Cov(\mu_{ijn}, \mu_{ijn-1}|s) &= Cov(\mu_{ijn-1}, \mu_{ijn-1} + \eta_{ijn}|s) \\ &= V(\mu_{ijn-1}|s) + Cov(\mu_{ijn-1}, \eta_{ijn}|s). \end{aligned}$$

It is the higher the more is known about the quality of the match. Next, consider the covariance of job-to-job movers. Recall that workers who are better matched at their old firm require a higher match quality at the new firm in order to switch jobs. Hence, although match qualities are independent across firms, i.e. $Cov(\mu_{ij'0}, \mu_{ijn-1}) = 0$, worker selection induces a positive correlation between the match quality at the old and new job, i.e. $Cov(\mu_{ij'0}, \mu_{ijn-1}|j) > 0$. Finally, consider job-to-unemployment movers. For these workers the autocovariance of the residual at time t and $t - 1$ is equal to the variance of ability, σ_f^2 . This is because workers have to search from scratch each

⁷Flinn (1986) derives a similar implication. He, however, does not distinguish between job-to-job and job-to-unemployment movers.

time they become unemployed. In other words, the optimal search strategy of an unemployed worker does not depend on the wage he earned at his previous job.

The variance of ability can therefore be recovered from the autocovariance of log-wage residuals for job-to-unemployment movers. This crucially depends on restrictions **R.1** (ability does not affect mobility) and **R.2** (workers lose the entire search capital). If workers do not lose the entire search capital when becoming unemployed, the match quality at the old job is positively correlated with the match quality at the new job also for job-to-unemployment movers, yielding an upward bias for the variance of ability. If ability affects the worker's mobility decision, the autocovariance for job-to-unemployment movers equals

$$Cov(u_{ijt}, u_{ijt-1} | \mathbf{u}) = V(f | \mathbf{u}) + Cov(f, \mu_n | \mathbf{u}) + Cov(f, \mu_0 | \mathbf{u}).$$

Although the distribution of ability is independent of the distribution of match quality, selection of workers into unemployment induces a correlation between the quality of the match and ability.

In order to get an idea how important these assumptions are for our results, we also compute the autocovariance of log-wage residuals for workers who left the firm for exogenous reasons. Following Gibbons and Katz (1991) and Dustmann and Meghir (2003), we use workers who are displaced from their firm because of plant closure as exogenous movers. This would relax assumption **R.1**, and yield an upward bias of the variance of ability if assumption **R.2** is violated. This is feasible with the German data, but not with the US data. Unfortunately, this is only feasible with the German data, but not with the US data.

Symmetric learning about ability Next, we discuss how symmetric learning about ability affects the autocovariance of log-wage residuals for job-to-unemployment movers. We maintain the assumption that ability does not affect mobility. We also assume that firms and workers separately learn about ability and match quality, i.e. they receive a separate signal for ability and match quality (as opposed to a signal for output only). If firms learn about ability, the perception of ability changes with time in the labour market. Hence, ability has to be indexed by t . The log-wage regression (4.6) changes

as follows.

$$\ln w_{ijt} = f_{it} + h(n) + 0.5S_n + \mu_{ijn} + \varepsilon_{ijt}.$$

Here, f_{it} denotes worker i 's expected ability at t . Bayesian updating implies that the best guess for the worker's ability tomorrow is his ability today. Hence, the covariance between f_t and f_{t-1} is equal to the variance of ability at $t-1$, $\sigma_{f_{t-1}}^2$. The autocovariance of log-wage residuals for job-to-unemployment movers at lag 1 therefore equals

$$\text{Cov}(u_{ijt}, u_{ijt-1} | \mathbf{u}) = \sigma_{f_{t-1}}^2.$$

Due to learning the variance of ability increases with experience, i.e. $\sigma_{f_t}^2$ increases with t . Hence, symmetric learning about ability implies that the autocovariance of log-wage residuals is higher for workers who become unemployed later in their career. The idea is similar to Farber and Gibbons (1996) who show that symmetric employer learning implies that wages follow a martingale. However, Farber and Gibbons do not distinguish between stayers, job-to-job and job-to-unemployment transitions.

The variance of log-wage residuals by experience

In this section we analyse how the variance of the log-wage residual varies with experience. Consider first workers who have just entered the labour market ($t = 0$). They start as unemployed and accept a wage offer if their log-wage corrected for ability, $\mu_{ij0} + 0.5S_0$, exceeds the reservation wage, $\Psi(0)$. The variance of log-wage residuals at labour market entry thus equals

$$V(u_{ij0}) = \sigma_f^2 + \sigma_\varepsilon^2 + V(\mu_{ij0} | \mu_{ij0} + 0.5S_0 \geq \Psi(0)). \quad (4.7)$$

The variance of the post-screening quality of the match at labour market entry, μ_0 , depends on the precision of the initial screen, τ_m , as well as on the precision of the prior quality of the match, τ_μ . From (4.3), μ_0 is normally distributed with variance $V(\mu_0) = \frac{\tau_m}{\tau_\mu(\tau_m + \tau_\mu)}$. Ignoring the impact of τ_m and τ_μ on the reservation wage, the variance of log-wage residuals is the higher the lower the precision of the prior match quality and the higher the precision of the screen.

Next, consider the variance of the residual of the second log-wage we observe for a worker ($t = 1$). Here, we have to distinguish between stayers, job-to-job and job-to-unemployment movers. The search problem of job-to-unemployment movers is the same as that of labour market entrants. Hence, for job-to-unemployment movers, the variance is equal to (4.7).

Suppose that for job-to-job movers the expected quality of the match at the new job is $\mu_{ij'0}$. Since tenure is equal to zero, the worker moves from job-to-job if the wage at the new job exceeds the wage at the old job, i.e. if $\mu_{ij'0} > \mu_{ij0}$. The variance of the log-wage residual for job-to-job movers therefore equals

$$V(u_{ij1}|j) = \sigma_f^2 + \sigma_\varepsilon^2 + V(\mu_{ij'0} | \mu_{ij'0} > \mu_{ij0}, \mu_{ij0} + 0.5S_0 \geq \Psi(0)).$$

Job-to-job movers are more selected than labour market entrants. Hence, for job-to-job movers the variance at $t = 1$, $V(u_{ij1}|j)$, is smaller than the variance at $t = 0$, $V(u_{ij0})$.

Finally, consider stayers. The following argument can also be found in Parent (2002). In appendix D we show that the unconditional distribution of the quality of the match at tenure $n = 1$ equals

$$V(\mu_1) = \frac{\tau_x + \tau_m}{\tau_\mu(\tau_\mu + \tau_x + \tau_m)}. \quad (4.8)$$

It is easy to check that $V(\mu_1) > V(\mu_0)$: As more information about the match becomes available, the variance of the quality of the match becomes closer to the variance of the true match quality, $1/\tau_\mu$. However, in the data we do not observe the unconditional variance, but the variance conditional on staying only. Suppose for simplicity that all workers have received one outside offer before the new information about the quality of the match arrived. The variance of the residual conditional on staying can then be written as

$$V(u_{ij1}|s) = \sigma_f^2 + \sigma_\varepsilon^2 + V(\mu_{ij1} | \mu_{ij0} > \mu_{ij'0}, \mu_{ij0} + 0.5S_0 \geq \Psi(0)).$$

It may be smaller or greater than the variance of log-wages at $t = 0$. On the one hand, the unconditional distribution of the expected quality of the match is becoming more dispersed as tenure increases. On the other hand, the group of workers who stay with

the employer is more selected than the group of workers who started the job, leading to a decrease in the variance. It is thus ambiguous whether the variance of the log-wage is higher at $t = 0$ or $t = 1$. More generally, the variance of the log-wage may increase or decrease with experience. There are two opposing effects. On the one hand, the revelation of new information about the quality of the match leads to an increase in the variance over time. The selection process, however, works in the opposite direction. A decline in the variance is evidence that workers allocate better matches as they become more experienced.

Symmetric learning about ability Finally, we discuss how these implications are affected if employers symmetrically learn about ability. Learning about ability implies that variance of ability increases with experience. Hence, the difference in the variance of log-wage residuals over time does not only reflect the change in the variance of match quality, but also the change in the variance of ability.

The covariance structure of the within-job wage growth residual

In a final step we analyse the covariance structure of the within-job wage growth residual. This provides evidence for the importance of learning about match quality. It also provides a way to identify the variance of the transitory wage component. Recall that the wage growth of stayers between time t and time $t - 1$ equals $\Delta \ln w_{ijt}|s = \Delta h(n) + 0.5\Delta S_n + \eta_{ijn}|s + \Delta \varepsilon_{ijt}$. Define $\Delta u_{ijt}|s$ as the residual of the within-job wage regression, i.e.

$$\Delta u_{ijt}|s = \eta_{ijn}|s + \Delta \varepsilon_{ijt}.$$

The variance and the covariance of Δu_{ijt} at lag l can be computed as

$$\begin{aligned} V(\Delta u_{ijt}|s) &= 2\sigma_\varepsilon^2 + V(\eta_{ijn}|s), \\ Cov(\Delta u_{ijt}, \Delta u_{ijt-1}|s) &= -\sigma_\varepsilon^2 + Cov(\eta_{ijn}, \eta_{ijn-1}|s), \text{ and} \\ Cov(\Delta u_{ijt}, \Delta u_{ijt-l}|s) &= Cov(\eta_{ijn}, \eta_{ijn-l}|s). \end{aligned}$$

The condition s indicates that the worker stayed with the same employer between t and $t - (l + 1)$ ⁸. The search model predicts a negative covariance at lag 1. This is because of the transitory wage component. At any higher lag the covariance should be close to zero. Importantly, the search model does *not* predict that workers whose wages have grown above average in the past will enjoy a wage growth above average in the future. This is different from a model in which workers differ with respect to the rate at which they accumulate firm-specific human capital. In appendix C we derive the covariance structure of the within-job wage growth residual in such a model. If the rate at which workers accumulate firm-specific human capital is heterogeneous, workers who experienced a high wage growth in the past will also experience a high wage growth in the future. Hence, such a model predicts a positive covariance at lag greater than 1.

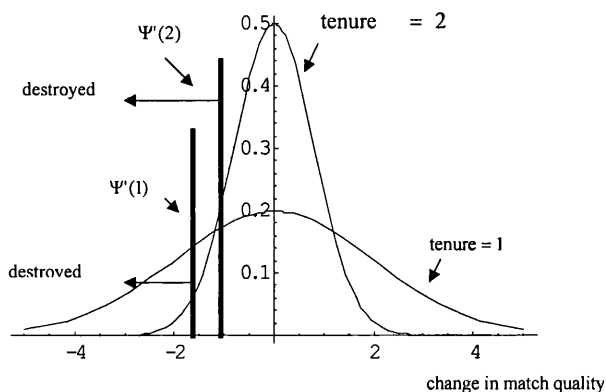
The variance of the within-job wage growth residual by tenure Next, we argue that the variance of the within-job wage growth residual declines with tenure if learning about match quality takes place. In appendix E we show that for a worker with tenure n , the unconditional distribution of the news about the quality of the match η_{ijn} is normal with mean 0 and variance

$$V(\eta_{ijn}) = \tau_{x\mu} S_n S_{n-1}. \quad (4.9)$$

Since the belief about the match quality becomes more precise with tenure, $V(\eta_{ijn})$ declines with tenure. In the data, however, we observe the variance of within-job wage growth conditional on staying only, i.e. we observe $2\sigma_\epsilon^2 + V(\eta_{ijn}|s)$. In the absence of firm-specific human capital accumulation the conditional variance still decreases with tenure if we also condition on the previous quality of the match, μ_{n-1} . To see this, consider a worker who has not received an outside offer between time t and $t - 1$. This worker stays with the employer if his new wage exceeds the reservation wage, i.e. if $S_n + \mu_{n-1} + \eta_n \geq \Psi(n)$. Define $\Psi'(n)$ as $\Psi(n) - (S_n + \mu_{n-1})$. A worker thus stays if $\eta_n \geq \Psi'(n)$. The variance of the within-job wage growth residual, conditional on the

⁸That is, the covariance at lag l is only observed in the data if the worker stayed with the same employer for at least $l + 1$ periods.

Figure 4.1: The variance of within-job wage growth residuals by tenure



previous quality of the match, thus equals

$$V(\Delta u_{ijt}|s, \mu_{ijn-1} = \mu) = 2\sigma_\varepsilon^2 + V(\eta_{ijn}|\mu_{ijn-1} = \mu, \eta_{ijn} > \Psi'(n)).$$

Since $\Psi'(n)$ is increasing in tenure because of learning about the quality of the match, this variance is decreasing in tenure⁹. This is illustrated by figure 4.1. The figure depicts the distribution of η_{ijn} at tenure equal to 1 and tenure equal to 2 for two workers who are equally well matched. At tenure 1 all jobs with $\eta_1 < \Psi'(1)$ are destroyed, while at tenure 2 all jobs with $\eta_2 < \Psi'(2)$ are destroyed. The variance of the change in the match quality conditional on staying is therefore higher at tenure 1 than at tenure 2.

The variance of the transitory wage component Next, we show how the variance of the transitory wage component can be recovered from the data. As tenure becomes large, there is nothing to learn about the quality of the match anymore. Hence, $V(\eta_{ijn})$ tends to zero and $\lim_{n \rightarrow \infty} V(\Delta u_{ijt}|\text{tenure} = n, s) = 2\sigma_\varepsilon^2$. The variance of the transitory wage component is thus identified from the variance of within-job wage growth for workers with high tenure. This crucially depends on restriction R.3, i.e. that the transitory wage component does not affect mobility. This clearly is the case if ε purely represents measurement error.

Symmetric learning about ability Finally, we discuss how learning about ability

⁹With firm-specific human capital accumulation $\Psi(n)$ is increasing in tenure, but $\Psi'(n)$ may not.

affects the variance of the within-job wage growth residual. If firms learn about ability, the within-job wage growth residual contains an additional term that captures the news in ability, i.e.

$$\Delta u_{ijt}|s = \Delta f_{it} + \eta_{ijn}|s + \Delta \varepsilon_{ijt}.$$

The variance now equals

$$V(\Delta u_{ijt}|s) = V(\Delta f_{it}) + 2\sigma_\varepsilon^2 + V(\eta_{ijn}|s).$$

As the belief of ability becomes increasingly precise with time in the labour market, $V(\Delta f_{it})$ decreases with experience. Hence, a decline in the variance of the within-job wage growth residual with tenure may reflect learning about ability rather than learning about match quality. The two types of learning differ as follows. Learning by ability is by experience; conditional on experience, tenure does not affect the belief about ability. Learning about match quality, in contrast, is by tenure; conditional on tenure, experience does not affect the belief about the quality of the match. Hence, a comparison of the variance by tenure for jobs that started at different points of time of the worker's career provides a clue about the importance of each type of learning.

4.3 Data description and variables

4.3.1 German data

The German data is of administrative nature, drawn from the German Social Security Records, the so-called IAB-Beschäftigtenstichprobe. It constitutes a 1 percent sample of the German labour force subject to social security contributions for the period between 1975 to 1995. The same data has been used in Chapter 3 for the comparison of the wage of workers who stay and leave the training firms. Details on the data, including the sample selection and variable definitions, can be found in appendix A. See also chapter 3, appendix B.

The data is well suited to analyse the process of job search. Most importantly, the data set contains the complete work history, including time spent in registered

unemployment and out of the labour force, from labour market entry onwards up to 20 years in the labour force. Furthermore, due to the administrative nature of the data, wages and employment spells are very accurate. Last but not least, wages can always be matched to a particular employer, and are never averaged across jobs. The major problem of the data is that it is right-censored at the highest level of earnings that are subject to social security contributions. Table 4.12 in appendix B tabulates the proportion of top-coded wage observations by experience and education. Top-coding is not a serious problem for unskilled workers and workers with an apprenticeship degree; for these workers less than 0.5 % of all wage observations are top-coded. For university graduates, however, the problem of top-coding is severe. Here, 10 % of all wage observations are top-coded. This proportion increases to 23 % for university graduates with ten years or more of experience. For this reason we restrict the empirical analysis that uses data on wages to unskilled workers and workers with an apprenticeship.

From this data base we select male individuals whom we observe from their entry in the labour market onwards since only for these workers experience and tenure can be measured precisely. We restrict analysis to former West Germany. See appendix A for details on the sample selection. Since the wage collection procedure changed in 1980, we drop observations before 1980 for the wage analysis, but not for the mobility analysis. Our final sample consists of 6122 unskilled workers, 25692 apprentices¹⁰, and 5203 university graduates, with a total of 355382 wage observations. Table 4.11 in appendix A reports means and standard deviations of the main variables used in the empirical analysis.

4.3.2 US data

The US data comes from the 1979 National Longitudinal Survey of Youth (NLSY79). At the time we started this analysis, the data was available from 1979 to 1996. The NLSY79 thus roughly covers the same time period as the IAB-Beschäftigtenstichprobe. As the German data the NLSY79 allows to construct a complete work history for

¹⁰The sample of apprentices is larger than in chapter 3 because we dropped the restriction that the worker has to complete the apprenticeship after 1980.

each respondent up to the date of the most recent interview. Furthermore, wages are uniquely assigned to particular jobs. As the US data is survey data whereas the German data is administrative data, measurement error is likely to be a more serious problem in the NLSY79.

The sample is the same as in Chapter 2. Details on the sample selection and variable description can be found in chapter 2, section 2.4 and appendix E. First, we consider only jobs after a transition from school to work has been made. Furthermore, we drop individuals who entered the labour market before 1978, as detailed information about weeks worked and employers worked for is only available from January 1978. We also delete all part-time jobs (< 30 hours a week), jobs without pay, and jobs of the self-employed. For multiple job holders, we define the main job as the job for which the worker worked most during the week. We only consider the main job for each respondent, and ignore multiple jobs. These selection criteria should minimise the problem that we overestimate mobility in the US because of jobs which are observed in the US data, but not in the German data, such as summer jobs. Our final sample consists of 550 high school dropouts, 2238 high school graduates and 871 university graduates with a total of 40860 wage observations. Results for high school graduates with and without some college education are very similar so that we decided not to distinguish between these two groups. Table 4.11 in appendix A reports means and standard deviations of the main variables used in the analysis.

One problem concerns the definition of unemployment in both data sets. We work with two definitions. The first definition is as follows. In the NLSY79 a respondent is considered as unemployed if he is actively looking for a job. If an interviewee does not report that he was actively looking for a job between two employment spells, the transition is considered as a job-to-job transition. In the IAB-Beschäftigtenstichprobe a transition is considered as a job-to-job transition if a worker has not been registered as unemployed between two employment spells. In Germany apprentices are eligible for unemployment benefits right after apprenticeship graduation. Unskilled workers and university graduates, in contrast, have to accumulate work experience before they become eligible for unemployment benefits. The probability that a worker who is

actively searching for a job, but not registered as unemployed, should therefore be small for apprentices, but may be substantial for unskilled workers and university graduates with little experience. For this reason we also use an alternative definition for a job-to-unemployment transition. According to this definition we classify a transition as a job-to-unemployment transition if the worker started the new job 4 weeks or later after his old job ended, independently of whether the worker was actively looking for a job or registered as unemployed.

4.4 Empirical evidence

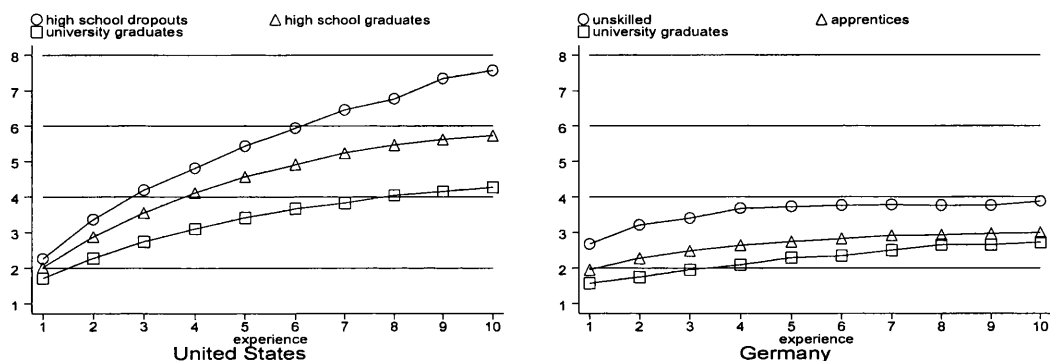
We first document differences in mobility rates between US and German workers (section 4.4.1). We then analyse the wage gains of stayers, job-to-job and job-to-unemployment movers (section 4.4.2). Finally, we analyse the covariance structure of wages and wage growth, and decompose the variance of the log-wage residual into the variance ability, a transitory component and match quality (section 4.4.3). This provides further evidence for the importance of match-specific productivity and job search.

4.4.1 Job-to-job and job-to-unemployment mobility

Number of jobs by experience Figure 4.2 demonstrates that mobility is substantially lower in Germany than in the US. The figure plots the average number of firms a worker has worked for by actual experience and education.

After ten years in the labour market high school graduates in the US are on their 6th job on average. This is in line with Topel and Ward (1992) who find that the typical American worker holds 7 jobs in ten years. In Germany, in contrast, workers who completed an apprenticeship have worked for only 2.8 employers in ten years. It is not only apprentices who are less mobile. There are also substantial differences between German unskilled workers and American high school graduates and particularly high school dropouts. Using the same data set, Dustmann and Meghir (2003) report very

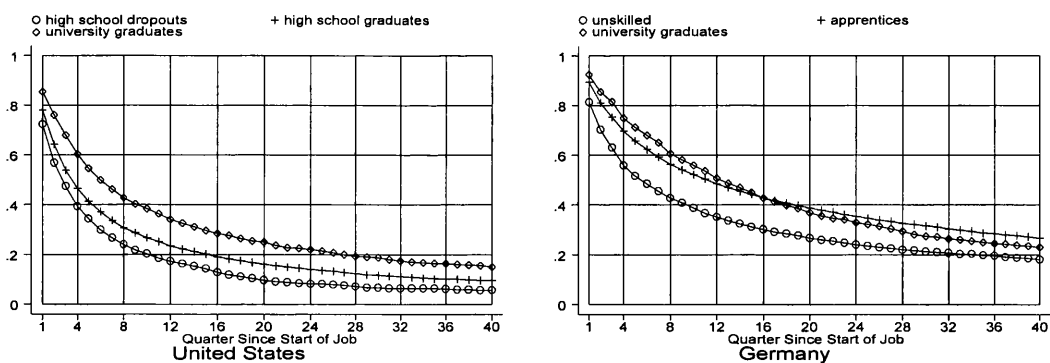
Figure 4.2: Number of jobs by experience and education



similar results¹¹. In both countries there is a negative correlation between mobility and education.

Job survival rates Although German workers are substantially less mobile than American workers, it is wrong to conclude that Germans do not change jobs. Figure 4.3 plots the empirical job survival rate for the three education groups.

Figure 4.3: Quarterly empirical job survival rate by education

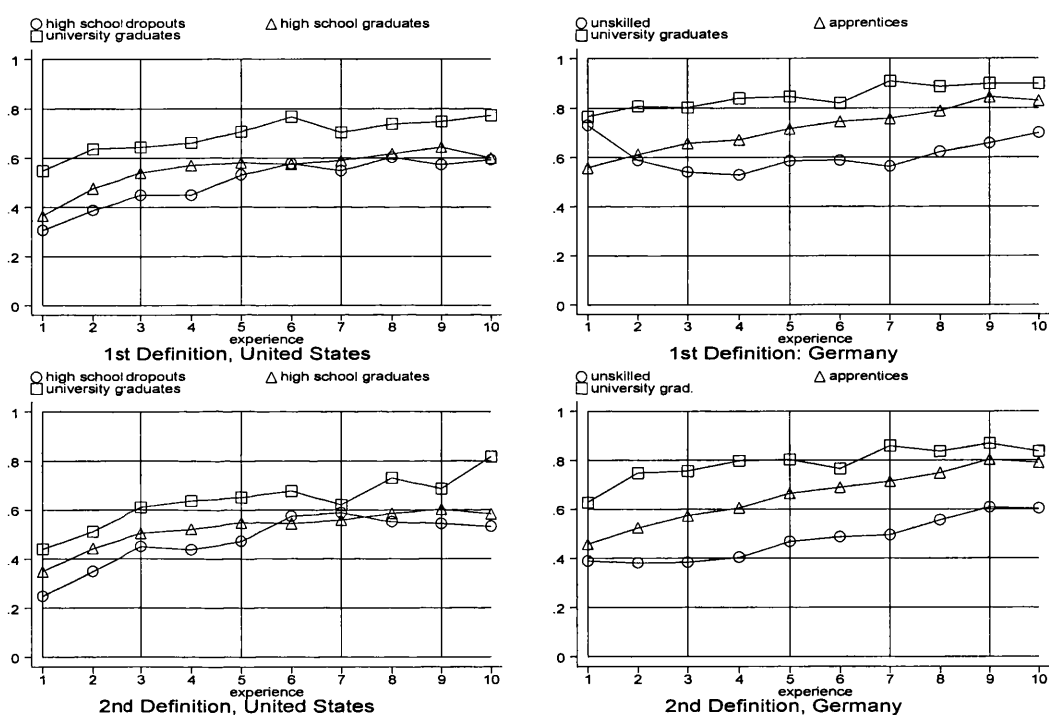


In Germany almost one third of jobs held by workers with an apprenticeship degree does not survive more than one year. This number increases to 45 % for unskilled workers. While these numbers seem quite high, they are low compared to those for the

¹¹Using data from the German Socio-economic Panel, Winkelmann (1996a, 1996b) also finds lower mobility rates for apprentices than for the unskilled.

United States. In the United States 54 % of the jobs held by high school graduates and 62 % of the jobs held by school dropouts terminate within the first year of employment.

Figure 4.4: The proportion of job-to-job transitions by experience: United States versus Germany



Job-to-job versus job-to-unemployment mobility The search model makes an important distinction between a job-to-job and a job-to-unemployment transition. If a worker moves from job-to-job, he gains search capital and allocates a better match. If, on the other hand, a worker becomes unemployed, he loses search capital and has to search from scratch. In order to get an idea about the relative importance of a job-to-job versus a job-to-unemployment switch in both countries, we plot the proportion of job-to-job transitions among all separations by actual experience (4.4). In the graph a worker is considered as a mover if he moved at least once in a given year of experience. The worker is classified as a job-to-job mover if he switched jobs at least once, but never became unemployed. The upper panel of figure 4.4 is based on the

first, while the lower panel is based on the second definition of a job-to-unemployment transition. The proportion of job-to-job transitions is higher in Germany than in the United States. This holds for all education groups, independently of which definition for a job-to-unemployment transition is used. There are also interesting similarities: In both countries the proportion of job-to-job transitions increases with experience and is higher for better educated workers. It is higher for the first than for the second definition of unemployment, especially at low levels of experience.

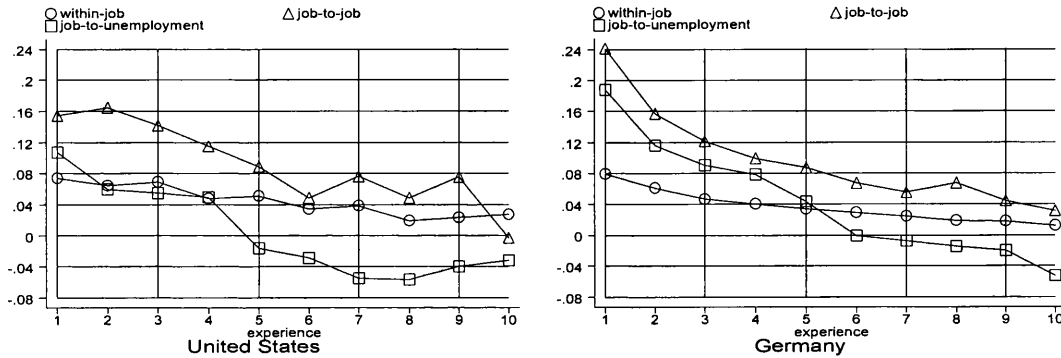
To sum up, mobility is substantially lower in Germany than in the US. It is not only apprentices who are less mobile, but also unskilled workers. German workers are particularly less likely to become unemployed than US workers.

4.4.2 Wage gains by mobility groups

In a next step, we analyse the gains and losses from staying, moving from job-to-job, and becoming unemployed. Since in the German data set many wage observations of university graduates are top-coded, we restrict the analysis to unskilled workers and workers who went through an apprenticeship only. Figure 4.5 plots the wage gains for stayers, job-to-job and job-to-unemployment movers by actual experience, pooled for all education groups. The definition for a stayer, job-to-job and job-to-unemployment mover is the same as in the previous section. The graph is based on the first definition of a job-to-unemployment transition. The choice of the definition has very little impact on the results.

Despite substantial differences in mobility rates, the wage growth of stayers is surprisingly similar in both countries. The wage gain of stayers is about 8 % in the first year. It then monotonically declines in both countries. The decline is somewhat more pronounced in Germany. The wage growth of job-to-job movers is considerably higher in Germany than in the US in the first year. In later years, however, it is remarkably similar in both countries. There are other striking similarities. In both countries the wage growth of job-to-job movers exceeds the wage growth of stayers. This difference declines with experience. This is consistent with the search model: As workers become better matched with experience, the gain in search capital from an additional job-to-job

Figure 4.5: Within-job, job-to-job and job-to-unemployment wage growth: United States versus Germany



transition typically declines with experience.

German and American workers mostly differ with respect to the wage gain (or loss) of job-to-unemployment movers. In the US job-to-unemployment movers roughly experience the same wage growth as stayers during the first four years in the labour market. At higher experience levels, the wage growth of job-to-unemployment movers is considerably lower than the wage growth of stayers. Experienced job-to-unemployment movers suffer a substantial wage loss. In Germany, we principally observe the same pattern. However, during the first five years the wage growth of job-to-unemployment movers exceeds that of stayers. Moreover, the wage loss at higher experience levels is less substantial in Germany than in the US¹².

4.4.3 The covariance structure of log-wage residuals

In this section we decompose the variance of the log wage residual into the variance of ability, match quality and a transitory component. This provides further evidence for the importance of match-specific productivity and gains from job search. We begin with the covariance structure of log-wage residuals by mobility groups and estimate the

¹²This pattern is consistent with the search model. Since it is the badly matched workers who are most likely to become unemployed, job-to-unemployment workers may on average be better matched at their new job than at their old job and thus gain in search capital. This effect should be strongest for workers with low experience.

variance of ability. We then analyse the covariance structure of the within-job wage growth residual and estimate the variance of the transitory wage component. Finally, we estimate the variance of the match quality at different experience levels.

The covariance structure of log-wage residuals by mobility groups

We begin with the autocovariance of the log-wage residual, $u_{ijt} = f_i + \mu_{ijn} + \varepsilon_{ijt}$, by mobility groups. First, we get an estimate for the residual. Let \hat{u}_{ijt} denote the estimated residual. We regress the log-wage on experience, experience squared, tenure, tenure squared, time dummies as well as race (US) and nationality (Germany), and obtain the residual.

Table 4.1: The autocovariance of log-wage residuals by mobility groups

	stayers	job-to-job movers	job-to-un. movers	exo. movers
unskilled, Germany	N=40706	N=5659	N=7355	N=1387
	0.1050	0.0495	0.0233	0.0345
apprentices, Germany	N=176514	N=22722	N=17822	N=4120
	0.0588	0.0393	0.0218	0.0315
high school dropouts, United States	N=2846	N=1227	N=1372	
	0.1075	0.0808	0.0571	
high school graduates, United States	N=14531	N=5139	N=4446	
	0.1282	0.0979	0.0716	
university graduates, United States	N=5558	N=1671	N=811	
	0.1607	0.1406	0.1055	

The table reports the autocovariance of log-wage residuals at lag 1 for stayers, job-to-job and job-to-unemployment movers. Results are based on the first definition of a job-to-unemployment transition. Exogenous movers are workers who lost their job due to plant closure. A worker is classified as an exogenous mover if he left the firm within the year the firm closed down or 1 year before that.

Table 4.1 shows the autocovariance of the estimated log-wage residual, \hat{u}_{ijt} , at lag 1 for three groups of workers: stayers, job-to-job and job-to-unemployment movers. The table 4.1 provides strong evidence that wages do not only depend on the worker's general, but also on his match-specific productivity. The pattern is remarkably similar

in both countries for all education groups: The autocovariance is highest for stayers, second highest for job-to-job movers, and smallest for job-to-unemployment movers¹³. This is implied by the search model. The autocovariance of residuals is higher for stayers than for movers since the quality of the match is independent across firms, but not within firms. The autocovariance is higher for job-to-job than for job-to-unemployment movers since workers who are well matched at their old job require a higher match at the new job in order to switch jobs. This induces a positive correlation between the quality of the match at the new and old job for job-to-job movers. Job-to-unemployment movers, in contrast, search from scratch and the match quality at the old job is uncorrelated with the match quality at the new job. Under the additional assumption that ability does not affect mobility, the autocovariance for job-to-unemployment movers identifies the variance of ability. In order to assess the importance of these assumptions, we also report the covariance for workers who left the firm because of plant closure. The idea here is that these workers left the firm for exogenous reasons. The autocovariance for exogenous movers is somewhat higher than that for job-to-unemployment movers, but lower than that for job-to-job movers. The autocovariance for job-to-unemployment movers may thus underestimate the variance of ability.

The variance of ability by experience

Next, we analyse how the autocovariance of job-to-unemployment movers changes with experience. If firms symmetrically learn about ability, the autocovariance of the log-wage residual is higher for workers who become unemployed later in their career.

Table 4.2 and table 4.3 display the autocovariance of log-wage residuals at lag 1 for movers who became unemployed at different points of time for the US and Germany, respectively. In the US, there is strong evidence for learning about ability for all education groups. For instance, for high school graduates the variance of ability increases

¹³Flinn (1986) also finds that the autocorrelation of log-wages is higher for those who stay than for those who move. He, however, does not distinguish between job-to-job and job-to-unemployment movers.

Table 4.2: The variance of ability by experience: United States

	high school dropouts	high school graduates	university graduates
≤ 1	N=278 0.0396	N=976 0.0446	N=224 0.0562
$1 < \text{exp} \leq 2$	N=261 0.0339	N=782 0.0506	N=129 0.0688
$2 < \text{exp} \leq 3$	N=184 0.0770	N=557 0.0755	N=114 0.0800
$3 < \text{exp} \leq 4$	N=161 0.0710	N=469 0.0898	N=93 0.1397
$4 < \text{exp} \leq 5$	N=107 0.0645	N=367 0.0892	N=55 0.1341
$5 < \text{exp} \leq 6$	N=100 0.0819	N=332 0.0855	N=49 0.1184
$6 < \text{exp} \leq 7$	N=73 0.0843	N=274 0.0957	N=37 0.1602
$7 < \text{exp} \leq 8$	N=69 0.1046	N=189 0.0948	N=34 0.1576
$\text{exp} > 8$	N=139 0.1068	N=500 0.0963	N=76 0.1683
total	N=1372 0.0571	N=4446 0.0716	N=811 0.1054

The table reports the autocovariance of log-wage residuals at lag 1 for movers who became unemployed at different points of time. Results are based on the first definition of a job-to-unemployment transition.

from 0.0446 for workers with no experience to 0.0963 for workers with more than 8 year of experience. The variance of ability is higher for better educated workers at all experience levels. For Germany, in contrast, we observe only a modest increase in the variance of ability¹⁴. The variance of ability is strikingly lower in Germany than in the US for all education groups at all experience levels.

¹⁴This is in line with existing tests for symmetric employer learning. Using a very different method to test for symmetric employer learning, Farber and Gibbons (1996) and Altonji and Pierret (2001) find strong evidence for symmetric learning in the US. Bauer and Haisken-DeNew (2001), on the other hand, find little evidence for symmetric employer learning in Germany.

Table 4.3: The variance of ability by experience: Germany

	unskilled	apprentices
≤ 1	N=2135 0.0221	N=7243 0.0197
$1 < \text{exp} \leq 2$	N=1427 0.0205	N=3560 0.0207
$2 < \text{exp} \leq 3$	N= 1143 0.0203	N= 2257 0.0208
$3 < \text{exp} \leq 4$	N=816 0.0226	N=1545 0.0237
$4 < \text{exp} \leq 5$	N=520 0.0261	N=1004 0.0269
$5 < \text{exp} \leq 6$	N=345 0.0262	N=705 0.0275
$6 < \text{exp} \leq 7$	N=288 0.0281	N=494 0.0284
$7 < \text{exp} \leq 8$	N=186 0.0274	N=353 0.0290
$\text{exp} > 8$	N=474 0.0275	N=661 0.0297
total	N=7355 0.0233	N=17822 0.0218

The table reports the autocovariance of log-wage residuals at lag 1 for movers who became unemployed at different points of time. Results are based on the first definition of a job-to-unemployment transition.

The covariance structure of within-job wage growth residuals

In this section we analyse the covariance structure of the within-job wage growth residual, $\Delta u_{ijt}|s$. This provides evidence for the importance of learning about match quality. It also provides a way to identify the variance of the transitory wage component. We begin the autocorrelation of the within-job wage growth residual. We then analyse how the variance of the within-job wage growth residual changes with tenure, and estimate the variance of the transitory wage component.

Autocorrelation of the within-job wage growth residual First, we get an estimate for the within-job wage growth residual. Let $\Delta\hat{u}_{ijt}|s$ denote this estimate. We regress the within-job wage growth on time dummies, the change in experience (and tenure), the change in experience squared as well as on the change in tenure squared, and obtain the residual¹⁵. Table 4.4 reports the autocorrelation of $\Delta\hat{u}_{ijt}|s$ at different time lags.

Table 4.4: Autocorrelation of the within-job wage growth residual

	lag 1	lag 2	lag 3
unskilled, Germany	N=28591	N=21653	N=16731
	-0.3020	-0.0015	-0.0052
apprentices, Germany	N=129345	N=97997	N=74936
	-0.3025	-0.0088	-0.0173
high school dropouts, United States	N=1522	N=967	N=642
	-0.3999	-0.0471	-0.0069
high school graduates, United States	N=9045	N=6246	N=4420
	-0.3018	-0.0067	-0.0242
university graduates, United States	N=3709	N=2595	N=1835
	-0.2339	-0.0256	-0.0302

The pattern is remarkably similar in both countries for all education groups. Without exception there is a strong negative autocorrelation at lag 1. At any higher lag the autocorrelation is close to zero and uniformly negative. Topel (1991) and Topel and Ward (1992) find a very similar pattern. This is the patterns that is implied by the search model: The negative correlation at lag 1 is driven by measurement error or transitory wage components. The very small and uniformly negative autocorrelation at higher lags implies that past wage growth is not a good indicator for future wage growth. Heterogeneity in predictable wage growth, as in a model with heterogeneous firm-specific human capital accumulation, is thus not an important feature of the data.

The variance of the within-job wage growth residual by tenure Next, we analyse how the variance of within-job wage growth residual, $\Delta\hat{u}_{ijt}|s$, varies with tenure.

¹⁵Race and nationality drops out of the within-job wage equation.

If learning about match quality is important, the variance should decrease with tenure as the belief about the quality of the match becomes increasingly precise with tenure.

Table 4.5: The variance of the within-job wage growth residual by tenure: United States

	high school dropouts			high school graduates			university graduates		
	all jobs	1st job	≥ 2	all jobs	1st job	≥ 2	all jobs	1st job	≥ 2
≤ 1	N=1130 0.0597	N=50 0.0688	N=630 0.0645	N=4706 0.0638	N=407 0.0755	N=2520 0.0583	N=1565 0.0896	N=205 0.1188	N=752 0.0928
$1 < \text{ten} \leq 2$	N=609 0.0597	N=40 0.0555	N=332 0.0570	N=2947 0.0511	N=296 0.0576	N=1590 0.0507	N=1143 0.0494	N=183 0.0942	N=544 0.0497
$2 < \text{ten} \leq 3$	N=351 0.0557	N=27 0.0339	N=183 0.0555	N=1887 0.0457	N=223 0.0437	N=1013 0.0464	N=777 0.0505	N=125 0.0567	N=335 0.0369
$3 < \text{ten} \leq 4$	N=234 0.0557	N=21 0.0408	N=116 0.0438	N=1331 0.0440	N=156 0.0378	N=690 0.0499	N=597 0.0399	N=108 0.0449	N=245 0.0414
$4 < \text{ten} \leq 5$	N=157 0.0489	N=18 0.0446	N=77 0.0495	N=1012 0.0345	N=124 0.0371	N=502 0.0303	N=401 0.0392	N=80 0.0393	N=138 0.0315
$5 < \text{ten} \leq 6$	N=95 0.0388	N=13 0.0320	N=44 0.0426	N=736 0.0393	N=105 0.0289	N=342 0.0526	N=314 0.0356	N=68 0.0369	N=95 0.0253
$6 < \text{ten} \leq 7$	N=70 0.0290	N=10 0.0312	N=27 0.0329	N=556 0.0387	N=75 0.0305	N=247 0.0403	N=248 0.0384	N=53 0.0323	N=66 0.0211
$7 < \text{ten} \leq 8$	N=60 0.0320	N=8 0.0446	N=22 0.0227	N=388 0.0294	N=63 0.0326	N=150 0.0309	N=181 0.0317	N=47 0.0245	N=44 0.0323
$\text{ten} \geq 8$	N=140 0.0381	N=21 0.0260	N=31 0.0228	N=967 0.0322	N=263 0.0313	N=213 0.0340	N=332 0.0378	N=117 0.0256	N=49 0.0262
all	N=2846 0.0566	N=208 0.0569	N=1462 0.0582	N=14531 0.0498	N=1714 0.0485	N=7267 0.0504	N=5558 0.0600	N=986 0.0674	N=2268 0.0596

For each education group, the first column reports the variance of the within-job wage growth residual, $V(\Delta\hat{u}_{ijt}|s)$, using all jobs. The second column reports the variance for the first job only. The third column reports the variance for jobs that started in the third year of experience or later (experience ≥ 2).

Table 4.5 and table 4.6 report the variance of $\Delta\hat{u}_{ijt}|s$ at different tenure levels for the United States and Germany, respectively. Column 1 reports results using all jobs. In line with the search model the variance tends to decline with tenure for all education groups in both countries. However, the decline may reflect learning about ability rather than learning about match quality. In order to get an idea about the importance of each type of learning, we report results separately for the first job (column 2) and jobs

that started in the third year in the labour market or later (experience ≥ 2 , column 3). A problem with this approach is that the sample size becomes very small, in particular for high school dropouts.

Table 4.6: The variance of the within-job wage growth residual by tenure: Germany

	unskilled			apprentices		
	all	1st job	≥ 2	all	1st job	≥ 2
ten ≤ 1	N=13218 0.0447	N=3906 0.0735	N=5184 0.0393	N=52850 0.0313	N=20525 0.0422	N=16786 0.0274
1 < ten ≤ 2	N=6754 0.0396	N=2281 0.0629	N=2684 0.0265	N=32420 0.0224	N=13688 0.0271	N=10068 0.0191
2 < ten ≤ 3	N=4842 0.0363	N=1761 0.0578	N=1858 0.0194	N=23336 0.0159	N=10398 0.0195	N=6871 0.0105
3 < ten ≤ 4	N=3612 0.0346	N=1410 0.0466	N=1337 0.0229	N=17448 0.0138	N=8235 0.0156	N=4717 0.0106
4 < ten ≤ 5	N=2813 0.0169	N=1181 0.0240	N=980 0.0115	N=11314 0.0143	N=6636 0.0154	N=3226 0.0120
5 < ten ≤ 6	N=2219 0.0166	N=964 0.0154	N=706 0.0178	N=10143 0.0121	N=5360 0.0138	N=2138 0.0099
6 < ten ≤ 7	N=1697 0.0110	N=803 0.0084	N=485 0.0111	N=7786 0.0113	N=4373 0.0124	N=1391 0.0100
7 < ten ≤ 8	N=1330 0.0121	N=685 0.0096	N=317 0.0154	N=6020 0.0096	N=3550 0.0109	N=889 0.0061
ten > 8	N=4241 0.0085	N=2557 0.0078	N=647 0.0077	N=13197 0.0090	N=8958 0.0087	N=1089 0.0090
total	N=40706 0.0327	N=15548 0.0416	N=14207 0.0234	N=176514 0.0203	N=81724 0.0235	N=47184 0.0168

For each education group, the first column reports the variance of the within-job wage growth residual, $V(\Delta\hat{u}_{ijt}|s)$, using all jobs. The second column reports the variance for the first job only. The third column reports the variance for jobs that started in the third year of experience or later (experience ≥ 2).

Consider first the results for the United States. In line with learning about ability, the variance of within-job wage growth is higher for the first job than for jobs that started later. From table 4.2, the variance of ability increases from 0.046 to 0.0502 for high school graduates in the first year in the labour market. The variance of within-job wage growth, using the first job only, declines somewhat more, leaving some scope for learning about match quality. However, overall the variance of ability increases by

more than the variance of within-job wage growth decreases. A similar picture emerges for university graduates and high school dropouts. Next, consider results for Germany. We again observe a lower variance for jobs that started later although, from table 4.3, learning about ability plays only a modest role in Germany. For both the first job and jobs that started during the second year of experience, the variance of within-job wage growth declines somewhat more than the variance about ability increases, for both unskilled workers and apprentices. In Germany there thus is evidence that some learning about match quality takes place.

Recovering the variance of the transitory wage component Next, we estimate the variance of the transitory wage component, σ_ε^2 . It is identified from the variance of the within-job wage growth residual for workers with high tenure, since $\lim_{n \rightarrow \infty} V(\Delta u_{ijt}|s) = 2\sigma_\varepsilon^2$: As tenure becomes large, nothing can be learned about the quality of the match anymore. The same argument holds for learning about ability. The variance of the within-job wage growth residual for jobs that lasted 8 years or longer is 0.0348 for the US, and 0.0085 for Germany, using all education groups. Hence, under the assumption that the variance of the transitory wage component is the same for all education groups and that all learning is completed within 8 years, our estimate for the variance of the measurement error is 0.0174 for the US data, and 0.00425 for the German data. The estimate for the US is somewhat lower than previous estimates¹⁶. The very low estimate for Germany confirms our guess that wages are measured very precisely due to the administrative nature of the data.

Unskilled versus apprentices A comparison between unskilled workers and apprentices in Germany turns out to be interesting. From table 4.5, the variance of the within-job wage growth residual is considerably higher for the unskilled than for apprentices at low tenure levels, particularly so at the first job. This difference all but disappears with tenure. One reason for the lower mobility of apprentices than for unskilled workers may thus be a higher precision of the quality of the match.

¹⁶For instance, Parent's (2002) estimates for the variance of transitory productivity shocks range from 0.0219 to 0.0427. Flinn's (1986) estimate is 0.031. Neither Parent nor Flinn take into account learning about ability.

To sum up, heterogeneity in the rate of firm-specific human capital accumulation is not an important feature of the data in either country. As expected, the variance of measurement error is considerably higher in the US data than in the German data. In Germany, the variance of the within-job wage growth residual is considerably higher for unskilled workers than for apprentices.

The variance of log-wage residuals by experience

In a final step we estimate the variance of match quality at different experience levels. The search model predicts that the variance of match quality eventually declines with experience as workers allocate better matches with time in the labour market. However, if learning about match quality is important, the variance may increase at low experience levels.

The first column of table 4.7 and 4.8 reports the variance of the log-wage residual at different experience levels for the US and Germany, respectively. In Germany we observe a gradual decline in the variance with experience for both education groups. In the US, on the other hand, the variance of log-wage residuals increases somewhat with experience. This is not surprising given the importance of learning about ability in the US. The second column of table 4.7 and 4.8 gives an idea about the variance of the match quality. It is obtained from the variance of the log-wage residual in column 1 by

- subtracting the variance of measurement error (i.e. 0.0174 for the US and 0.00425 for Germany), and
- subtracting the variance of ability from table 4.2 and 4.3, respectively.

We now observe a decline in the variance of the quality of the match for all education groups in both countries. Most of the decline takes place during the early years in the labour market when workers are most mobile and gains from moving from job-to-job are highest. This again provides evidence for the importance of match-specific productivity and search capital in both countries.

Table 4.7: The variance of the log-wage residual and match quality by experience:
United States

	high school dropouts		high school graduates		university graduates	
	1	2	1	2	1	2
0	N=550		N=2238		N=871	
	0.1276	0.0706	0.1408	0.0788	0.2191	0.1455
exp ≤ 1	N=698		N=2844		N=958	
	0.1216	0.0703	0.1416	0.0736	0.1984	0.1122
1 < exp ≤ 2	N=737		N=2974		N=1017	
	0.1396	0.0452	0.1476	0.0547	0.2081	0.1107
2 < exp ≤ 3	N= 660		N= 2641		N=939	
	0.1300	0.0416	0.1544	0.0472	0.2112	0.0539
3 < exp ≤ 4	N=587		N=2473		N=912	
	0.1451	0.0532	0.1557	0.0491	0.1845	0.0330
4 < exp ≤ 5	N=498		N=2216		N=796	
	0.1444	0.0451	0.1565	0.0536	0.1901	0.0543
5 < exp ≤ 6	N=467		N=2046		N=732	
	0.1519	0.0502	0.1625	0.0494	0.2080	0.0304
6 < exp ≤ 7	N=422		N=1833		N=655	
	0.1416	0.0196	0.1656	0.0533	0.1955	0.0205
7 < exp ≤ 8	N=358		N=11626		N=581	
	0.1441	0.0199	0.1583	0.0446	0.2098	0.0241
exp > 8	N=1018		N=5463		N=1450	
	0.1440	0.0198	0.1511	0.0374	0.2182	0.0325

For each education group, the first column reports the variance of the log-wage residual. The second column reports our estimate for the variance of the quality of the match. It is obtained from the variance in column 1 by subtracting the variance of measurement error (0.0174), and the variance of ability from table 4.2.

A comparison between unskilled workers and apprentices in Germany is particularly interesting. The variance of the quality of the match is almost three times as high for unskilled workers than for apprentices at low levels of experience. This difference all but disappears with experience. Furthermore, there is little evidence that the variance of the quality of the match is higher in the US than in Germany. The variance as well its decline with experience is highest for unskilled workers in Germany. It is, despite substantial differences in mobility rates, of remarkably similar magnitude for German apprentices and American high school graduates and dropouts.

Table 4.8: The variance of the log-wage residual and match quality by experience: Germany

	unskilled		apprentices	
	1	2	1	2
0	N=6122		N=25962	
	0.2199	0.1935	0.0965	0.0725
≤ 1	N=10898		N=41278	
	0.1717	0.1469	0.0844	0.0595
$1 < \text{exp} \leq 2$	N=7920		N=32248	
	0.1497	0.1251	0.0774	0.0523
$2 < \text{exp} \leq 3$	N= 6652		N= 27398	
	0.1238	0.0969	0.0714	0.0434
$3 < \text{exp} \leq 4$	N=5715		N=23501	
	0.0987	0.0683	0.0702	0.0391
$4 < \text{exp} \leq 5$	N=4749		N=20125	
	0.0964	0.0659	0.0667	0.0349
$5 < \text{exp} \leq 6$	N=4000		N=17078	
	0.0849	0.0525	0.0678	0.0351
$6 < \text{exp} \leq 7$	N=3249		14649	
	0.0835	0.0518	0.0647	0.0314
$7 < \text{exp} \leq 8$	N=2931		12213	
	0.0738	0.0421	0.0662	0.0322
$\text{exp} > 8$	N=11336		N=32911	
	0.0645	0.0327	0.0610	0.0271

For each education group, the first column reports the variance of the log-wage residual. The second column reports our estimate for the variance of the quality of the match. It is obtained from the variance in column 1 by subtracting the variance of measurement error (0.00425), and the variance of ability from table 4.3.

Table 4.9 and table 4.10 show the proportion of the variance of the residual that can be attributed to ability and match quality, for the US and Germany. In both countries and for all education groups, the proportion of the variance attributable to ability increases with experience, while that attributable to match quality declines with experience. The proportion attributable to match quality is considerably higher in Germany than in the US. This may indicate that search frictions play a more important role in Germany.

To sum up, the variance of the quality of the match declines with experience in

Table 4.9: The variance of ability and match quality by experience: United States

	high school dropouts		high school graduates		university graduates	
	ability	match quality	ability	match quality	ability	match quality
exp=0	31.0 %	55.4 %	31.7 %	55.9 %	25.6 %	66.4 %
2 < exp ≤ 3	54.6 %	32.0 %	58.2 %	30.6 %	66.1 %	25.5 %
4 < exp ≤ 5	53.9 %	29.7 %	52.6 %	32.9 %	62.3 %	28.5 %
exp > 8	74.1 %	13.7 %	63.7 %	24.7 %	77.1 %	14.9 %

The proportion of the variance attributable to ability and match quality is reported. The estimate for the variance of measurement error is 0.0174. Estimates are based on table 4.2 and 4.8.

Table 4.10: The variance of ability and match quality by experience: Germany

	unskilled		apprentices	
	ability	match quality	ability	match quality
exp=0	10.0 %	87.9 %	20.4 %	75.1 %
2 < exp ≤ 3	18.2 %	78.3 %	33.2 %	60.8 %
4 < exp ≤ 5	27.1 %	68.4 %	40.6 %	51.5 %
exp > 8	42.6 %	50.7 %	48.7 %	44.4 %

The proportion of the variance attributable to ability and match quality is reported. The estimate for the variance of measurement error is 0.00425. Estimates are based on table 4.3 and 4.8.

both countries for all education groups. The proportion of the variance that can be attributed to match quality is considerably higher in Germany. In Germany the variance as well as its decline with experience is considerably stronger for unskilled workers than for apprentices.

4.5 Summary

This chapter compares job mobility of young men in two countries with very different labour market institutions, the United States and Germany. The empirical analysis for Germany is based on a new and unique administrative data set, the so-called IAB-Beschäftigtenstichprobe. The data for the US comes from the 1979 National Lon-

gitudinal Survey of Youth (NLSY79). Both data sets are well suited to analyse the mobility behaviour of young workers. Most importantly, both data sets allow to construct an accurate and complete work history of workers from labour market entry onwards, including all job-to-unemployment transitions.

We set up a search model that guides the empirical analysis. The search model borrows from Jovanovic (1984) and Mortensen (1988). The two crucial features of the model are on-the-job search and learning about match quality. The job thus is (partly) an experience good. On-the-job search leads to endogenous job-to-job transitions and allows workers to allocate better matches over time. Learning about the quality of the match, on the other hand, leads to endogenous job-to-unemployment transitions: A match that used to be preferable to unemployment may not be preferable anymore after disappointing news about the match arrived. A distinction between job-to-job and job-to-unemployment transitions is important. If a worker moves from job-to-job he gains search capital, whereas he loses search capital and has to search from scratch if he becomes unemployed.

We first derive the restrictions the search model imposes on wage gains of stayers, job-to-job and job-to-unemployment movers, and propose a simple method of decomposing the variance of the log-wage residual into the variance of ability, match quality and a transitory component.

Our main findings can be summarised as follows. Mobility is substantially lower in Germany than in the US for all education groups. German workers are particularly less likely to become unemployed. Yet, there are important similarities.

- Match-specific productivity plays an important role in both countries for all education groups. As predicted by the search model, the autocovariance of log-wage residuals is highest for stayers, second highest for job-to-job movers and lowest for job-to-unemployment movers. Furthermore, the proportion of the variance that can be attributed to match declines with experience, while the proportion that can be attributed to ability increases with experience in both countries for all education groups.

- There is some evidence for learning about match quality in both countries, as the variance of the within-job wage growth residual declines with tenure.
- Heterogeneity in the rate of firm-specific human capital accumulation is not an important feature of the data in either country, as the autocovariance of the within-job wage growth residual is close to zero and uniformly negative at lag greater than 1.

American and German workers mostly differ with respect to the variance of ability and transitory wage component.

- The variance of ability is considerably higher in US for all education groups. Furthermore, learning about ability plays an important role in the US, but only a modest role in Germany. The proportion of the variance attributable to match quality is considerably higher in Germany than in the US. This may indicate that search frictions play a more important role in Germany.
- As expected, the variance of the transitory wage component is substantially lower in the German administrative data than in the US survey data due to measurement error.

There are also interesting differences between unskilled workers and apprentices in Germany. First, the variance of the quality of the match is almost three times as high for unskilled workers than for apprentices at low levels of experience. This difference all but disappears with experience. This is consistent with the idea that job search plays a more important role for unskilled workers than for apprentices. Second, the variance of the within-job wage growth residual is considerably higher for unskilled workers than for apprentices at low tenure levels, particularly so at the first job. This difference all but disappears with tenure. One reason for the lower mobility of apprentices than for unskilled workers may thus be a higher precision of the quality of the match.

4.6 Appendix

A Data

Table 4.11: Means and standard deviations of selected variables: US and German data

	NLSY79			IAB		
	hs drop.	hs grad.	uni. grad.	unskilled	appr.	uni. grad.
number of workers	550	2238	871	6122	25962	5203
proportion	15.03 %	61.61 %	23.80 %	16.41 %	69.63 %	13.95 %
total number of observations	5995	26354	8911	63752	247363	44257
US: hourly wage rate in \$	8.47	10.04	15.25	112.40	133.62	198.25
GER: daily wage rate in DM	(8.7092)	(8.4914)	(9.5673)	(40.36)	(38.90)	(49.35)
experience	5.09	5.53	5.22	4.81	4.54	5.11
	(3.6494)	(3.8742)	(3.5779)	(4.1295)	(3.5402)	(4.2876)
tenure	1.87	2.29	2.71	2.71	2.93	3.22
	(2.2942)	(2.7087)	(2.6828)	(3.2938)	(2.9588)	(3.3429)
age	25.46	26.21	28.06	23.82	25.25	31.00
	(4.6164)	(4.5769)	(4.1986)	(4.4851)	(3.8448)	(5.0814)
number of jobs	5.33	4.60	3.47	3.13	2.35	2.03
	(3.4782)	(3.0828)	(2.2722)	(2.7676)	(1.8934)	(1.3929)

Sample Selection: German data We include all apprentices in the sample that are also included in the sample of apprentices in chapter 3. See Chapter 3, appendix B for the selection criteria. We drop the restriction that the worker has to complete the apprenticeship after 1979. Unskilled workers and university graduates are included in the sample if

- no A-level (*Abitur*), polytechnic or university degree at labour market entry: not older than 15 in 1975, and at most 19 at labour market entry;
- A-levels (*Abitur*), no polytechnic or university degree at labour market entry: not older than 19 in 1975, at most 21 at labour market entry;

- polytechnic degree: not older than 23 in 1975, at most 27 at labour market entry;
- university degree: not older than 24 in 1975, at most 29 at labour market entry.

Definition of education: German data

1. See Chapter 3, appendix B, for a definition of an *apprentice*.
2. A *university graduate* is a worker who held at least one job that classified him as a university or polytechnic graduate.
3. An *unskilled* worker is a worker who is neither classified as an apprentice or university graduate.

B Censoring in the German data

Table 4.12 shows the proportion of censored wages by education and experience. For university graduates, censoring is severe. By the 10th year of experience, however, almost every fourth wage is top-coded.

Table 4.12: Wage censoring by experience and education: German data

experience	unskilled	apprentices	university graduates
0	0.11 %	0.19 %	2.44 %
1	0.11 %	0.25 %	5.33 %
2	0.26 %	0.25 %	7.21 %
3	0.21 %	0.40 %	9.09 %
4	0.33 %	0.62 %	12.54 %
5	0.30 %	0.58 %	15.10 %
6	0.29%	0.71 %	17.93 %
7	0.53 %	0.66 %	21.93 %
8	0.43 %	0.80 %	23.80 %
9	0.36 %	0.88 %	23.02 %
Total	0.22 %	0.44 %	10.11 %

C The covariance structure of the within-job wage growth residual if the rate of firm-specific human capital accumulation is heterogeneous

In this section we derive the covariance structure of the within-job wage growth residual if more able workers accumulate more firm-specific human capital. Suppose for simplicity that firm-specific human capital accumulation is linear, i.e. $h_i(n) = \beta_i n$. Let $\bar{\beta}$ denote the average return to firm-specific human capital accumulation. Further suppose for simplicity that $t - (t - 1) = 1$ for all wage informations. In order to focus on the implications of heterogeneous firm-specific human capital accumulation, we assume that the match is a pure inspection good. The wage growth of stayers then equals $\ln_{ijt} | s = \bar{\beta} + (\beta_i - \bar{\beta})$. The variance and the covariance of within-job wage growth at lag l can be computed as

$$\begin{aligned} V(\Delta w_{ijt} | \text{tenure} = n, s) &= 2\sigma_\epsilon^2 + V(\beta_i | s), \\ \text{Cov}(\Delta w_{ijt}, \Delta w_{ijt-1} | \text{tenure} = n, s) &= -\sigma_\epsilon^2 + V(\beta_i | s) \leq 0, \text{ and} \\ \text{Cov}(\Delta w_{ijt}, \Delta w_{ijt-l} | \text{tenure} = n, s) &= V(\beta_i | s) \geq 0. \end{aligned}$$

The condition s indicates that the worker stayed with the same employer between t and $t - (l + 1)$. A model with heterogenous firm-specific human capital accumulation therefore implies a positive covariance at lag greater than 1.

D The distribution of the updated quality of the match

From (4.1),

$$\mu_0 = \frac{\bar{\mu}\tau_\mu + m\tau_m}{\tau_\mu + \tau_m} = \bar{\mu} + \frac{\tau_m(m - \bar{\mu})}{\tau_\mu + \tau_m}.$$

Hence, $E[\mu_0] = \bar{\mu}$. $V(\mu_0)$ equals

$$V(\mu_0) = E \left[\left\{ \frac{\tau_m(m - \bar{\mu})}{\tau_\mu + \tau_m} \right\}^2 \right] = \frac{\tau_m^2(1/\tau_\mu + 1/\tau_m)}{(\tau_\mu + \tau_m)^2} = \frac{\tau_m}{\tau_\mu(\tau_\mu + \tau_m)}.$$

This is expression (4.3). From (4.4),

$$\mu_n = \frac{\bar{\mu}\tau_\mu + m\tau_m + n\tau_x\bar{x}}{\tau_\mu + \tau_m + n\tau_x} = \bar{\mu} + \frac{\tau_m(m - \bar{\mu}) + n\tau_x(\bar{x} - \bar{\mu})}{\tau_\mu + \tau_m + n\tau_x}.$$

Hence, $E[\mu_n] = \bar{\mu}$. $V(\mu_n)$ can be computed as

$$\begin{aligned} V(\mu_n) &= E \left[\left\{ \frac{\tau_m(m - \bar{\mu}) + n\tau_x(\bar{x} - \bar{\mu})}{\tau_\mu + \tau_m + n\tau_x} \right\}^2 \right] \\ &= \frac{\tau_m^2(1/\tau_\mu + 1/\tau_m) + n^2\tau_x^2(1/\tau_\mu + 1/n\tau_x) + 2n\tau_m\tau_x/\tau_\mu}{(\tau_\mu + \tau_m + n\tau_x)^2} \\ &= \frac{\tau_m + n\tau_x}{\tau_\mu(\tau_\mu + \tau_m + n\tau_x)}. \end{aligned}$$

Setting $n = 1$ gives expression (4.8).

E The distribution of the innovation in the quality of the match

Note that $\mu_n = \frac{S_n}{S_{n-1}}\mu_{n-1} + S_n\tau_x x_n$, where x_n is the n th signal observation on-the-job. Consequently,

$$\mu_n - \mu_{n-1} = S_n\tau_x(x_n - \mu_{n-1}).$$

Hence, $E[\mu_n - \mu_{n-1}] = 0$. $V(\mu_n - \mu_{n-1})$ can be computed as

$$V(\mu_n - \mu_{n-1}) = S_n^2\tau_x(V(x_n) + V(\mu_{n-1}) - 2Cov(\mu_n, \mu_{n-1})).$$

Using (4.8) and (4.5),

$$\begin{aligned} V(\mu_n - \mu_{n-1}) &= S_n^2\tau_x(1/\tau_x + 1/\tau_\mu + S_{n-1}\frac{\tau_m + (n-1)\tau_x}{\tau_\mu} - 2S_{n-1}\frac{\tau_m + (n-1)\tau_x}{\tau_\mu}) \\ &= \frac{S_{n-1}S_n^2\tau_x}{\tau_x\tau_\mu}(\tau_\mu(n\tau_x + \tau_m + \tau_\mu)) = S_nS_{n-1}\tau_x. \end{aligned}$$

This is expression (4.9).

Chapter 5

Concluding Remarks

There are three benchmark models, or 'building blocks', for explaining labour market outcomes: human capital models, search and matching models, and learning models. The objective of this thesis is to assess the empirical importance of each building block model.

Chapter 2 focuses on learning models. The question addressed in this chapter is: Is employer learning symmetric or asymmetric? A model that nests both learning hypotheses is developed. It is shown that existing tests that have been derived under one learning hypothesis are also consistent with the other. New tests that allow to discriminate between the two hypotheses are proposed. The chapter also contributes to the literature on job mobility by analysing the impact of ability and learning about ability on the worker's decision to switch jobs. Evidence from the NLSY79 indicates that asymmetric employer learning plays an important role for university graduates, but not for high school graduates. The way the market acquires new information about workers' productivity thus differs across education groups. This has important implications. First, the results imply that labour markets work more efficiently for high school dropouts and high school graduates than for university graduates. Second, the results suggest that firms earn -ex post- higher rents on university graduates than on high school dropouts and graduates. Third, they imply that firms have a stronger incentive to sponsor general training for university graduates than for high school

graduates and dropouts. This may be one explanation for Lynch's (1992) finding that in the US better educated workers are more likely to receive company-related training.

Chapter 3 focuses on human capital models. It revisits the question why firms pay for training even if training is general. The question is addressed within the German apprenticeship system. The focus is on the impact of wage rigidities caused by unions on training. A model of firm-sponsored training and unions is developed and empirical implications are derived. The empirical evidence suggests that wage floors created by unions are an important reason for firm-financed training in Germany. However, asymmetric employer learning cannot be ruled out as an additional reason for firm-financed training.

These results have important policy implications. First, a deregulation of the German labour market that limits the power of unions may have undesired consequences for the apprenticeship system. Second, attempts to introduce training schemes in countries in which unions play only a minor role may fail. It is, however, important to bear in mind that the reason why unions induce firms to sponsor training is that unions make long-term wage contracts enforceable. There may be ways to do so other than through the unionisation of the economy.

The results also have important welfare implications. Most importantly, there is no evidence that unions reduce training in the economy. Moreover, the results are compatible with unions moving training closer to the socially optimal level. This crucially depends on the assumption that firms cannot fully commit to training provision. If instead firms can fully commit to training provision, our results imply that in Germany unions raise apprenticeship training above the socially optimal level. Distinguishing between the limited and full commitment case is an important task for future research.

Chapter 4 focuses on search and matching models. It compares job mobility of young men in two countries with very different labour market institutions, the United States and Germany. Based on a search model with endogenous job-to-job and job-to-unemployment transitions, I propose a simple method of decomposing the variance of wages into an ability component, a match quality component and a transitory com-

ponent. Match-specific productivity and job search plays an important role in both countries, for all education groups. The proportion of the variance that is attributable to match quality is considerably higher in Germany than in the US. In Germany, there is evidence that job search and match-specific productivity matters more for unskilled workers than for apprentices.

Let me conclude by pointing out the difference chapter 2 and chapter 4 have taken to job mobility. Chapter 4 ignores the effect of ability, and focuses on the impact of match-specific productivity as well as learning about match-specific productivity on job mobility. Chapter 2, in contrast, ignores the effect of match-specific productivity, and focuses on the impact of ability as well as learning about ability on job mobility. Developing a model that integrates these two aspects of job mobility and is at the same time tractable is an important direction for future research.

Bibliography

- [1] Abraham, K.G. and H.S. Farber (1987), "Job Duration, Seniority, and Earnings", *American Economic Review*, 77, pp. 278-297.
- [2] Acemoglu, D. and J.-S. Pischke (1998), "Why Do Firms Train: Theory and Evidence", *Quarterly Journal of Economics*, 113, pp. 79-119.
- [3] Acemoglu, D. and S. Pischke (1999a), "Beyond Becker: Training in Imperfect Labour Markets", *Economic Journal Features*, 109, pp. F112-F142.
- [4] Acemoglu, D. and S. Pischke (1999b), "The Structure of Wages and Investment in General Training", *Journal of Political Economy*, 107, pp. 539-573.
- [5] Acemoglu, D. and S. Pischke (2001), "Minimum Wages and On-the-Job Training", *IZA Discussion Paper No. 384*.
- [6] Altonji, J.G. and C.R. Pierret (2001), "Employer Learning and Statistical Discrimination", *Quarterly Journal of Economics*, 116, pp. 313-350.
- [7] Altonji, J.G. and R.A. Shakotko (1987), "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54, pp. 437-459.
- [8] Altonji, J.G. and J.R. Spetzler (1991), "Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training", *Industrial and Labour Relations Review*, 45, 58-71.
- [9] Altonji, J.G. and N. Williams (1997), "Do Wages Rise with Seniority: A Reassessment", *NBER Working Paper 6010*.

- [10] Altonji, J.G. and N. Williams (1998), "The Effects of labour Market Experience, Job Seniority, and Job Mobility on Wage Growth", in: Polachek, S.W. (ed.), *Research in labour economics*, 17, London: JAI Press, pp. 233-276.
- [11] Autor, D. (2001), "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics*, 116, pp. 1409-1448.
- [12] Baker, G., Gibbs, M. and B. Holmstrom (1994a), "The Internal Economics of the Firm: Evidence from Personnel Data", *Quarterly Journal of Economics*, 109, pp. 881-919.
- [13] Baker, G., Gibbs, M. and B. Holmstrom (1994b), "The Wage Policy of a Firm", *Quarterly Journal of Economics*, 109, pp. 921-955.
- [14] von Bardeleben, R., Beicht, U. and Feher, K. (1995), "Betriebliche Kosten und Nutzen der Ausbildung. Repräsentative Ergebnisse aus Industrie, Handel und Handwerk", *Berichte zur beruflichen Bildung*, Heft 187, Bielefeld: Bertelsmann.
- [15] Bauer, T.K. and J.P. Haisken-DeNew (2001), "Employer Learning and the Returns to Schooling", *Labour Economics*, 8, pp. 161-180.
- [16] Becker, G. (1964), "Human Capital", Chicago: The University of Chicago Press.
- [17] Bellmann, L., Kohaut, S. and C. Schnabel (1998), "Ausmaß und Entwicklung der übertariflichen Entlohnung" *IW-Trends*, 25, pp. 5-14.
- [18] Berkovitch, E. (1990), "A Stigma Theory of Unemployment Duration", in: Weiss, Y. (ed.), *Advances in the Theory and Measurement of Unemployment*. New York: St. Martin's Press, pp. 20-56.
- [19] Bernhardt, D. (1995), "Strategic Promotion and Compensation", *Review of Economic Studies*, 62, pp. 315-339.
- [20] Bernhardt, D. and D. Scoones (1993), "Promotion, Turnover, and Preemptive Wage Offers", *American Economic Review*, 84, pp. 771-791.

- [21] Bhaskar, V. and S. Holden (2002), "Wage Differentiation vs Generalised Training", mimeo, University of Essex.
- [22] Boeheim, R. and A. Booth (2002), "The Impact of Bargaining Institutions on Employer-Provided Training in Britain", mimeo, University of Essex.
- [23] Booth, A., and Chatterji, M. (1998), "Unions and Efficient Training", *Economic Journal*, 108, 328-343.
- [24] Booth, A., Francesconi, M., and G. Zoega (1999), "Training, Rent-Sharing and Unions", Working Papers of the Institute for Labour Research, Paper 99/23, Colchester, University of Essex.
- [25] Burdett, K. (1978), "A Theory of Employee Job Search and Quit Rates", *American Economic Review*, 68, pp. 212-220.
- [26] Carneiro, P., J. Heckman, and E. Vytlačil (2003), "Estimating Returns to Education when it Varies among Individuals", University of Chicago, mimeo.
- [27] Card, D. (1992), "The Effect of Unions on the Distribution of Wages: Redistribution or Relabelling? NBER Working Paper No 4195.
- [28] Cawley, J., Heckman, J., and E. Vytlačil (2001), "Three Observations on Wages and Measured Cognitive Ability", *Labour Economics*, 8, pp. 419-442.
- [29] DeGroot, M.H. (1970), "Optimal Statistical Decisions", New York: McCraw-Hill.
- [30] Dertouzos, J. and L. Karoly (1992), "labour-Market Responses to Employer Liability", RAND, Santa Monica.
- [31] Dinardo, J., Fortin, N. and T. Lemieux (1996), "labour Market Institutions and the Distribution of Wages: A Semi-Parametric Approach, *Econometrica*, 64, pp. 1001-1044.
- [32] Doiron, D.J.(1995), "Lay-Offs as Signals: The Canadian Evidence", *Canadian Journal of Economics*, 28, pp. 899-913.

- [33] Dustmann, C. and C. Meghir (2003), "Wages, Experience, and Seniority", IFS Working Paper W01/01.
- [34] Farber, H.S. (1994), "The Analysis of Interfirm Worker Mobility", *Journal of labour Economics*, 12, pp. 554-593.
- [35] Farber, H.S. (1999), "Mobility and Stability: The Dynamics of Job Change in labour Markets", in: Ashenfelter, O. and D. Card (eds.), *Handbook of labour Economics*, Volume 3b, New York: Elsevier Science, North-Holland, pp. 459-485.
- [36] Farber, H.S. and R. Gibbons (1996), "Learning and Wage Dynamics", *Quarterly Journal of Economics*, 111, pp. 1007-1047.
- [37] Farber, H. and A. Krueger (1992), "Union Membership in the United States: the Decline Continues", NBER Working Paper # 4216.
- [38] Flinn, C.J. (1986), "Wages and Job Mobility of Young Workers", *Journal of Political Economy*, Part 2, 94, pp. S88-S110.
- [39] Franz, W. and D. Soskice (1995), "The German Apprenticeship System", in: F. Buttler, W. Franz, R. Schettkat, and D. Soskice (eds.), *Institutional frameworks and labour market performance*, London and New York: Routledge, pp. 208-304.
- [40] Freeman, R.B., and J. L. Medoff (1984), "What Do Unions Do?" , New York: Basic Books.
- [41] Fortin, N. and T. Lemieux (1997), "Institutional Change and Wage Inequality: Is there a Linkage?", *Journal of Economic Perspectives*, 11, pp. 75-96.
- [42] Fudenberg, D. and J. Tirole (2000), "Customer Poaching and Brand Switching", *RAND Journal of Economics*, 31, pp. 634-657.
- [43] Galindo-Rueda, F. (2002), "Employer Learning and Schooling-Related Statistical Discrimination in Britain", mimeo, University College London.
- [44] Gibbons, R. and L.F. Katz (1991), "Layoffs and Lemons", *Journal of labour Economics*, 9, pp. 351-380.

- [45] Green, F., S. Machin, D. Wilkinson (1996), "Trade Unions and Training Practices in British Workplaces", *Industrial and labour Relations Review*, 52, pp. 179-195.
- [46] Greenwald, B.C. (1986), "Adverse Selection in the labour Market", *Review of Economic Studies*, 53, pp. 325-347.
- [47] Grund, C. (1999), "Stigma Effects of Layoffs? Evidence from German Micro-Data", *Economics Letters*, 64, pp. 241-247.
- [48] Harhoff, D. and T.J. Kane (1997), "Is the German Apprenticeship System a Panacea for the U.S. labour Market?" *Journal of Population Economics*, 10, pp. 171-196.
- [49] Harris, M. and B. Holstrom (1982), "A Theory of Wage Dynamics", *Review of Economic Studies*, 49, pp. 315-333.
- [50] Hashimoto, M. (1981), "Human Capital as a Shared Investment", *American Economic Review*, 71, 153-161.
- [51] Heckman, J., and E. Vytlačil (2001), "Identifying the Role of Cognitive Ability in Explaining the Level of Change in the Return to Schooling", *Review of Economics and Statistics*, 83, pp. 1-12.
- [52] Jovanovic, B. (1979a), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, 87, pp. 1246-1260.
- [53] Jovanovic, B. (1979b), "Firm-specific Human Capital and Turnover", *Journal of Political Economy*, Part 1, 87, pp. 972-990.
- [54] Jovanovic, B. (1984), "Matching, Turnover, and Unemployment", *Journal of Political Economy*, 92, pp. 108-122.
- [55] Kölling, Arndt (2000), "The IAB-Establishment Panel", *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 120, pp. 291-300.
- [56] Krashinsky, H. (2002), "Evidence on Adverse Selection and Establishment Size in the labour Market", *Industrial and labour Relations Review*, 56, pp. 84-96.

- [57] Kugler, A.D. and G. Saint-Paul (2000), "Hiring and Firing Costs, Adverse Selection, and Long-Term Unemployment", Universitat Pompeu Fabra, Economics and Business Working Paper 447.
- [58] Lange, F. (2004), "The Returns to Schooling and Ability During the Early Career: Evidence on Job Market Signaling", University of Chicago, mimeo.
- [59] Lazear, E. (1984), "Raids and Offer-Matching", NBER Working Paper 1419.
- [60] Lehmann, Karen (2001), "Veränderung und Stabilität der Flächentarifbindung von Arbeitgebern in Deutschland: Eine theoretische und empirische Analyse", PhD-thesis, University of Hannover.
- [61] Lemieux, T. (1998), "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Nonrandom Selection", *Journal of Labour Economics*, 16, pp. 261-291.
- [62] Light, A. and K. McGarry (1998), "Job Change Patterns and the Wages of Young Men", *Review of Economics and Statistics*, 80, pp. 276-286.
- [63] Loewenstein, M.A. and J.R. Spetzler (1997), "Delayed Formal Training", *Industrial and Labour Relations Review*, 51, 82-99.
- [64] Loewenstein, M.A. and J.R. Spetzler (1999), "General and Specific Training: Evidence and Implications", *Journal of Human Resources*, 18, 710-733.
- [65] Lynch, L. M. (1992), "Private-Sector Training and Earnings of Young Workers", *American Economic Review*, 82, pp. 299-312.
- [66] Meghir, C. and M. Palme (2001), "The Effect of a Social Experiment in Education", IFS Working Paper WP01/11.
- [67] Montgomery, J.D. (1999), "Adverse Selection and Employment Cycles", *Journal of Labour Economics*, 17, pp. 281-297.
- [68] Mortensen, D.T. (1988), "Wages, Separations, and Job Tenure: On-the-Job Specific Training or Matching?", 6, pp. 445-471.

- [69] Mortensen, D.T. and C.A. Pissarides (1994), "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies*, 61, pp. 397-415.
- [70] Neal, D. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers", *Journal of Labour Economics*, 13, pp. 653-677.
- [71] Neal, D. (1998), "The Link between Ability and Specialization", *Journal of Human Resources*, 33, pp. 173-200.
- [72] Neal, D. (1999), "The Complexity of Job Mobility among Young Men", *Journal of Labour Economics*, 17, pp. 237-261.
- [73] Nelson, P. (1970), "Information and Consumer Behavior", *Journal of Political Economy*, 78, pp. 311-329.
- [74] Neumark, D. and W. Wascher (2001), "Minimum Wages and Training Revisited", *Journal of Labour Economics*, 19, pp. 563-595.
- [75] OECD (1999), *Economic Outlook*.
- [76] Oettinger, G.S. (1996), "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap", *Journal of Labour Economics*, 14, pp. 52-78.
- [77] Parent, D. (2000), "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics", *Journal of Labour Economics*, 18, pp. 306-323.
- [78] Parent, D. (2002), "Matching, Human Capital, and the Covariance Structure of Earnings", *Labour Economics*, 9, pp. 375-404.
- [79] Postel-Vinay, F. and J.M. Robin (2002), "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70, pp. 2295-2350.
- [80] Rodriguez-Planas, N. (1998), "Playing Hard to Get: Theory and Evidence on Layoffs, Recalls, and Unemployment", Boston University, Industry Studies programme Working Paper Series on Economics, 86.

- [81] Ryan, P. (2001), "The School-to-Work Transition: A Cross-National Perspective", *Journal of Economic Literature*, 39, pp. 34-92.
- [82] Shi, S. (2002), "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology", *Review of Economic Studies*, 69, pp. 467-491.
- [83] Soskice, D. (1994), "Reconciling Markets and Institutions: The German Apprenticeship System", in: L.M. Lynch (ed.) *Training and the Private Sector, International Comparisons*, NBER Series in Comparative labour Markets, Chicago: University of Chicago Press, pp. 1-24.
- [84] Topel, R.H. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Seniority", *Journal of Political Economy*, 99, pp. 145-176.
- [85] Topel, R.H. and M.P. Ward (1992), "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics*, 107, pp. 439-479.
- [86] Waldmann, M. (1984), "Job Assignment, Signalling, and Efficiency", *Rand Journal of Economics*, 15, pp. 255-267.
- [87] Waldman, M. (1996), "Asymmetric Learning and the Wage/Productivity Relationship", *Journal of Economic Behavior and Organization*, 31, pp. 419-429.
- [88] Winkelmann, R. (1996a), "Training, Earnings, and Mobility in Germany", *Konjunkturpolitik*, 42, pp. 275-298.
- [89] Winkelmann, R. (1996b), "Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany", *Industrial and labour Relations Review*, 49, pp. 658-672.
- [90] Winkelmann, R. (1997), "How Do Young Workers Get Their Training: A Survey of Germany vs the US", *Journal of Population Economics*, 10, pp. 159-170.