Transitions on the Labor Market: Unemployment, Transfer Receipt and the Low-Wage-Sector

D I S S E R T A T I O N

zur Erlangung des akademischen Grades
doctor rerum politicarum

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
der Freien Universität Berlin

von

Dipl.-Volksw. Arne Uhlendorff,
geboren am 06. Januar 1974 in Ratzeburg

eingereicht im Februar 2007
Tag der Disputation: 05.06.2007

Erstgutachter: Prof. Dr. Klaus F. Zimmermann

Zweitgutachter: Prof. Irwin Collier, Ph.D.

Preface

I have written this dissertational thesis while being a research associate at the German Institute for Economic Research. During the past four years I have received support by my supervisor, colleagues and friends. First of all I would like to thank my supervisor Klaus F. Zimmermann for his guidance and his constant encouragement and support for my research.

My research profited very much from the discussions with and the support of my colleagues and friends Katharina Wrohlich and Peter Haan. The cooperation with them was central for a very friendly and motivating working atmosphere.

I thank my colleagues at the DIW Berlin for their help and support, among others Karl Brenke, Tilman Brück, Marco Caliendo, Joachim Frick, Johannes Geyer, Peter Krause, Michal Myck, Thorsten Schneider and Jürgen Schupp. Moreover I would like to thank Ingo Geishecker, Stephen Jenkins, Hilmar Schneider and Christian Stoltenberg for helpful comments on and discussions about my work. My research has profited from the cooperation with several co-authors. I thank Hilmar Schneider (Chapter 2), Klaus F. Zimmermann (Chapter 3) and Peter Haan (Appendix), whose joint papers are part of this dissertation.

In addition to that I thank the German Science Foundation (DFG) for financial support through SPP1169 "Potential for more Flexibility on Heterogeneous Labor Markets", seminar participants at several conferences and the participants of the Berlin Labour Research Network (BeNA) for comments on my work.

Contents

Chapter 1

Introduction

Unemployment in Germany has increased regularly over the last decades. While the unemployment rate goes up in recessions, it does not recover to the same degree in booms. To put it differently, unemployment is no longer a cyclical phenomenon but a structural problem. In 2005, around five million persons or 13.0% of the civilian labor force were registered as unemployed (Statistisches Bundesamt, 2006). Reduction of this high unemployment rate is one of the most important and challenging issues the German society is confronted with.

The risk of unemployment is especially high among lows-skilled and unskilled individuals and the correlation of the level of education and the risk of unemployment has increased in recent years. In 2004, around one fifth of low-skilled workers were unemployed, compared to merely one out of twenty in the 1970s (Reinberg and Hummel, 2005). This gap in the risk of unemployment between highly educated workers and workers with low education is relatively large in Germany by international standards (OECD, 2006). Moreover, since the early 1970s, the unemployment rates of natives and migrants in Germany diverge. In 2005, the average share of unemployed migrants has been 25.2% in comparison to the much lower 11.9% among natives (Bundesagentur für Arbeit, 2006). This may be partly explained by the education and skill level of foreign nationals in Germany, which is rather low, but could also result from additional, ethnic-specific disadvantages on the labor market.

It is often argued that the establishment of a low-wage sector is necessary to overcome the high unemployment rate among low skilled workers, and that the current tax and transfer system induces low incentives to work, especially for the low skilled (Zimmermann, 2003). There have been various proposals to increase work incentives in the recent German policy debate. These included a reduction in social assistance benefits and introduction of workfare or in-work benefits. For a discussion of several reform-proposals, see e.g. Steiner (2004) or Bonin and Schneider (2006).

This study is a contribution to the ongoing debate about the determinants of individual employment dynamics. In particular, I focus on the probability of entering and leaving unemployment and receipt of social assistance. Why do some households leave social assistance for a job and others don't? Do migrants stay longer unemployed or are they shorter employed? Does there exist a "low pay - no pay" cycle in Germany? These questions are analyzed in this book. In Chapter 2, I study the duration of social assistance and its incentive effects on the probability of leaving welfare in favor of paid work. Chapter 3 focusses on differences between natives and migrants with respect to unemployment duration and subsequent employment stability. In Chapter 4, I analyze the mobility between three labor market states over time: working in low paid jobs, working in higher paid jobs and not working.

All analysis are based on the German Socio Economic Panel Study (SOEP), a representative longitudinal study of private households which was started in 1984. Specifically, it provides detailed information on employment states and earnings of all household members as well as information on their migration history. The longitudinal character of the SOEP allows to study individual labor market dynamics described by the duration of different states and the transition probabilities between them, and to control for unobserved heterogeneity, which is an important issue for the estimation of duration and transition processes.

Contribution of this Study

It is often argued that the high level of assistance claims in Germany induces little incentive for workers with low productivity to seek for a job. In Chapter 2, I examine the influence of the ratio between estimated potential labor income and the assistance payment level on the probability of leaving social assistance.[1] The potential net labor income is estimated with standard wage equations by accounting for sample selection and applying a simple tax function. Estimating a discrete time hazard rate model with competing risks and unobserved heterogeneity, the results show that the ratio has a positive effect on the probability of leaving social assistance. This effect is especially relevant for households with a potential labor income higher than their assistance payment level. The result is different from previous studies dealing with the determinants of social assistance duration in Germany but in line with international evidence. The difference derives from a simultaneous consideration of both sources of income, the net household labor income and the social assistance level, and additionally from a differentiation between transitions to work and alternative transitions. This is important because potential labor income is only relevant for transitions into work.

Another reason of concern is that unemployment is typically very high among migrants with a tendency to rise over time. This higher rate of unemployment could result from a higher risk of becoming unemployed, i.e. higher frequency of unemployment spells or shorter periods of employment, as well as from a lower probability of leaving unemployment, i.e. longer duration of unemployment spells. Hence, in Chapter 3, I investigate both sources of higher unemployment rates, unemployment duration and subsequent employment stability.[2] The two processes are determined by observed and unobserved characteristics and it is reasonable to assume that the unobserved characteristics influencing both processes are not independent from each other. Therefore, I estimate unemployment and subsequent employment duration models simultaneously and allow for correlation between unobserved terms. The results show that, compared to natives with the same observable and unobservable characteristics, unemployed migrants do not find less stable positions but they need more time to find these jobs. The probability of leaving unemployment also varies strongly between eth-

[1] Chapter 2 is based on joint work with Hilmar Schneider (Schneider and Uhlendorff, 2004). The main results of the study are reported in Schneider and Uhlendorff (2005).

[2] Chapter 3 is based on joint work with Klaus F. Zimmermann (Uhlendorff and Zimmermann, 2006).

nicities. The first and second generation of migrants from Turkey are identified as the major risk group.

Unskilled and low-skilled individuals have a relatively high risk of unemployment. In this context low pay employment is evaluated differently. On the one hand, it is argued that rising employment rates in the low pay sector could be one solution to overcome the high unemployment rate among low-skilled workers. On the other hand, low paid jobs are often associated with unstable working careers and high risk of unemployment. Therefore, it is important to know whether low paid jobs are transitory experiences of the working career and stepping stones to better jobs or whether there exists a "low pay - no pay" cycle. In Chapter 4, I analyze low pay and non-employment dynamics of men in west Germany. The focus lies on the extent of true or genuine state dependence in low pay and non-employment. I estimate dynamic multinomial logit panel data models with random effects taking the initial conditions problem and potential endogeneity of panel attrition into account. In line with results from other countries, this first study on Germany finds true state dependence in low pay jobs and confirms previous results of state dependence in non-employment. Moreover, I find evidence for a "low pay - no pay" cycle, i.e. being low paid or not employed itself increases the probability of being in one of these states in the next year. However, compared to not working, being low paid does not have adverse effects on future employment prospects: the employment probability increases with low pay employment and the probability of being high paid seems to be higher for previously low paid workers relative to not working. The Appendix contains a description of the applied simulation procedure in the context of random effects multinomial logit models.[3]

In Chapter 5, I summarize the main findings of the empirical analyses and derive conclusions. I also discuss potential shortcomings and problems of the analyses and provide an outlook for further research in this area.

[3]The Appendix is based on joint work with Peter Haan (Haan and Uhlendorff, 2006).

Conclusions and Policy Implications

The policy conclusions which can be drawn from the results of these analyses are the following. For the social assistance the results show that higher benefit levels lead to longer social assistance spells. A reduction of the benefit level could be one solution to overcome the incentive problems. However, the amount of social assistance is related to a basic minimum income concept and a general reduction of the benefit level would be unlikely to find political support. A reduction of the social assistance level is not the only way to overcome incentive problems of a transfer program, there exist other possible solutions, for example workfare. If the payment of benefits is made conditional on the willingness to accept regular employment or a workfare job, the attractiveness of low-paid jobs in the labor market relative to benefit receipt would increase.

The analysis of the unemployment dynamics among migrants and natives identifies migrants from Turkey as the major risk group among the migrants. Therefore, policy should concentrate on the job finding process of Turkish migrants to fight their disadvantages on the labor market. The focus should be primarily on the probability of finding a job and to a lower degree on the quality of the job, because once Turkish migrants find a job, this is as stable as the jobs of natives and other migrants. However, more research is needed to understand *why* Turkish migrants stay longer unemployed than natives.

Low paid jobs are often associated with unstable working careers and high risk of unemployment. However, I find evidence that low paid jobs are stepping stones to better jobs and no evidence that being low paid does have any adverse effects on future employment prospects if it is compared to non-employment. Therefore, policy should try to increase employment rates in the low wage sector. From a methodological point of view, the study shows that it is important to control for the endogeneity of the initial state in low pay dynamics. Moreover, I do not find evidence for endogeneity of panel attrition. This result is important for further practical applications, because the simpler model without potential endogeneity of panel attrition goes along with a substantial reduction in computation time.

Chapter 2

The Transition from Welfare to Work and the Role of Potential Labor Income[1]

2.1 Introduction

In Germany, the number of welfare recipients as well as the amount of income support expenditures have been rising almost continuously in the past. In the year 2002, about 2.8 million persons in 1.4 million households received social assistance and the expenditures amount to 25 billion euro.[2] The share of the municipalities' revenues spent for the permanent social welfare transfers has been rising from 3.5% in 1980 to 6.8% in 2002 (Haustein and Krieger, 2004). What is the reason for this large number of welfare recipients? In the economic literature as well as in public debate on the German welfare system the incentive argument plays an important role. It asserts that if the difference between the level of transfers and potential income from a regular job is too small then picking up a job is not attractive for the individual (see for example Ochel, 2003). In this paper, we analyze this hypothesis by estimating the impact of the ratio between potential labor income and the amount of transfer payment on the transition probability from welfare to employment in Germany.

[1]The following analysis is based on joint work with Hilmar Schneider (Schneider and Uhlendorff, 2004).

[2]This number of welfare recipients refers to permanent transfers, the so-called *Hilfe zum Lebensunterhalt*, described in detail in section 2. The amount of expenditures refer to permanent transfers and transfers for persons in special circumstances.

The German social assistance is a means-tested transfer program financed by the municipalities. The receipt of transfer payments requires that the household income including other transfer payments like unemployment benefits does not exceed a certain minimum level. In contrast to the unemployment benefits, everybody is principally eligible, irrespective of his or her individual employment history. Although the receipt is in principle unlimited, only a minority of households stays on welfare over a longer period of time. Assistance claims expire as soon as alternative income exceeds a certain threshold. This may be due to labor income but could also be due to changes in household formation or the receipt of alternative transfer payments like pensions etc.

Numerous studies exist on the duration of income support spells. Most of them are referring to North America and dealing with women receiving welfare. A typical result says that the probability of leaving welfare is higher for better educated and white persons and declines with the number of (young) children, disabilities, the amount of benefits and the level of regional unemployment (see e.g. Blank, 1989, Stewart and Dooley, 1999, or Gittleman, 2001, a summary is given by Moffitt, 1992). For the U.S. find for female heads of households that a decreasing amount of welfare benefits reduces the welfare dependency. Barrett (2000) finds a positive effect of the educational attainment on the welfare exit rate, having a greater impact for women than for men. With a higher educational attainment he assumes a higher offered wage and therefore a higher relative attractiveness of employment. Fortin, Lacroix, and Drolet (2004) use information from a natural experiment in 1989 in the Province of Quebec in Canada. They conclude that a doubling of the amount of social assistance for single men and women aged under 30 significantly increased their individual spell duration. Lemieux and Milligan (2006) make use of a sharp discontinuity (at the age of 30) in the level of social assistance in Quebec to identify the effect of the benefit level on employment and find strong evidence that more generous social assistance benefit levels reduce employment.

Hazard rate models are an appropriate tool for the analysis of the duration of welfare receipt. For Germany duration analyses of social assistance usually do not take income variables into account (see for example Voges and Rohwer, 1992, Gangl, 1998, or Gebauer, Petschauer, and Vobruba, 2002). Gangl (1998) has shown that it is important to distinguish transitions to employment from alternative transitions like transitions out of the labor market. In descriptive analyses, the social assistance levels are generally compared with the average wage of a special group of employees, for example unskilled workers in manufacturing (e.g. Engels, 2001, or Boss, 2002). From these descriptive statistics conclusions about the incentives for work for different household types are drawn, but these hypotheses are not tested econometrically. As far as we know, there exist only two studies testing the influence of income variables on the duration of welfare receipt. The study by Riphahn (1999) on basis of the German Socio-Economic Panel study (SOEP) shows no significant influence of a predicted real net income variable for full-time employed individuals on the exit probability out of income support. However, she does not take the amount of social transfers into account. Wilde (2003) examines the difference between social benefits and the average income for unskilled employees on the probability of leaving social welfare using the Low Income Panel and finds no significant effects. Both, Wilde and Riphahn do not distinguish between different transitions in their regression analysis.

In our analysis we use data from the SOEP. Between 1992 and 2000 retrospective monthly information about social welfare receipt for each month of the previous calendar year is part of the household questionnaire. Spell duration is observed in months, but generated by a continuous time process. Taking into account the discrete time measurement of the underlying data, we estimate a discrete-time proportional hazard rate model with competing risks and risk specific unobserved heterogeneity. We assume that the destination specific hazard rates are constant within each interval and allow for dependent competing risks via a correlation of the random intercepts.

Controlling for several typical covariates the ratio between potential labor income and the welfare level shows a positive effect on the probability of leaving social welfare for work. This effect is especially relevant for households with a potential labor income higher than their social welfare level. In contrast to previous studies, we cannot reject the incentive hypothesis for Germany. The alternative hypothesis, that the higher probability of transition is a consequence of a higher rate of job offers for better educated persons, seems to be of minor relevance: The effect of the ratio keeps significant when controlling for education and the local labor market performance.

Section 2.2 of this chapter gives a short description of the system of social welfare in Germany and its theoretical implications on labor supply. Section 2.3 provides information on the data and the estimated models. Section 2.4 presents empirical results and section 2.5 concludes.

2.2 Incentive Effects of Social Assistance in Germany

The German social assistance (*Sozialhilfe*) is a means-tested transfer program and consists of two main parts: Permanent transfers to households with low income (*Hilfe zum Lebensunterhalt*, *HLU*) and transfers to persons in special circumstances who need temporary financial support[3] (*Hilfe in besonderen Lebenslagen*). In this study we concentrate on the HLU because these payments are principally unlimited and may act as a permanent alternative to a labor income. In the following, the terms welfare and social assistance are used as synonyms and refer to HLU. The receipt of social assistance requires that the household income including other transfer payments like unemployment benefits (*Arbeitslosengeld* and *Arbeitslosenhilfe*, the latter is also means-tested and principally unlimited) does not exceed a certain minimum level.

In principle, everybody in need may claim for social assistance, while unemployment benefits are only accessible to those who have previously contributed to unemployment insurance for a minimum period within a given time frame. Moreover, the amount of unemployment benefits depends on the income in the previous job, while the amount of social assistance is related to a basic minimum income concept depending on household size and household composition. In addi-

[3] For example, pregnant women or homeless persons searching for a new apartment.

tion, the eligibility criteria in case of own income differ between the means-tested unemployment benefits and the social assistance. Therefore, the analysis is restricted to social assistance spells without taking into account spells of the also means-tested *Arbeitslosenhilfe*.[4]

Welfare benefits consist of basic allowances for every adult household member, housing allowances and one-time payments. The amount for basic allowances differs between the federal states depending on the regional minimal costs of living. In 2003, it ranged between 282 and 297 euro per month. Children get 50-90 percent depending on age. Expectant mothers, older and disabled persons receive higher basic allowances than "normal" adults. In principle, the amount of social assistance fills the gap between own income and the maximum benefit for the household. Labor income up to 25% of the basic allowance is not taken into account. Additional income is deducted at an implicit marginal tax rate of 85% until the deduction exceeds 50% of the basic allowance. Above this threshold the implicit marginal tax rate is 100%.

The impact of social assistance on work incentives can be described in terms of a very basic utility model for the choice between consumption and leisure (see for example Blundell and MaCurdy, 1999, or Moffitt, 2002). Assume a utility maximizing individual subject to a non convex budget set. A stylised depiction is given in Figure 2.1.

If a person is not working at all, he or she will receive social assistance at a level of V. If this person works at his or her net market wage rate of w up to Q hours per week, disposable net income will not increase, since earned income is totally deducted from social assistance. Only when the number of hours worked is exceeding Q, disposable income will increase with slope w. The resulting non convex budget set is expressed in the graph by line B.

If no social assistance existed, it would be optimal to work H hours per week with a disposable income of wH. Beside the budget set, optimal labor supply is a function of individual preferences that are responsible for the shape of the

[4]Recipients of unemployment benefits are included in the analysis if they are members of households receiving social assistance.

Figure 2.1: The impact of social assistance on work incentives for a stylized budget set

grey indifference curve I_1 tangent to the budget line. The utility level associated with H hours of work per week has to be compared to the utility level resulting from not working at all, which is expressed by the dotted indifference curve I_0. In the depicted case, not working generates a higher utility level than working. A utility maximizing person would only work H hours per week, if he or she would be able to achieve a disposable income of at least Y^*, which can also be expressed in terms of an implicit minimum wage rate. Note that Y^* in the example given is more than twice as high as V.

From this simple static perspective, the individual has perfect information about jobs and is faced with a particular wage. In this framework, periods of welfare receipt and subsequent employment can not be explained. In contrast to that, in a dynamic job-search model an individual is not faced with a particular wage but with a particular distribution of wages (e.g. Devine and Kiefer, 1991). To leave a welfare program for employment requires an acceptable job offer. The exit rate from welfare to work λ depends on the arrival rate of job offers ω and on the job acceptance rate θ. It can be written as a product of both: $\lambda(t) = \omega(t)\theta(t)$

Wage offers are only accepted if they exceed the reservation wage. This reservation wage depends positively on the amount of social benefits. Given that wage offers arrive at a certain frequency and given a level of market (or expected) wage of an individual, it is more likely to observe exits from social assistance, the lower the welfare payment. The effect of the level of market wage is theoretically ambiguous. On the one hand, the reservation wage should increase with the expected wage. On the other hand the expected value of a job is higher which should increase the job search intensity. We assume that the positive effects of an increase in the expected wage on the probability of leaving social welfare outweigh the negative effects for the group of welfare recipients in our study. Furthermore we assume that the effect of the difference between the two income sources depends on the relative level of the social benefits. Therefore, households with a lower ratio between potential labor income and the amount of social welfare should have a lower hazard rate from welfare to employment.[5] This is the hypothesis we are going to test in the empirical part of the paper.

In our context, not single persons but households receive transfer payments and one has to decide how to calculate the potential household income. We calculate this income variable assuming one adult household member working full-time (for similar approaches see Riphahn, 1999, or Wilde, 2003).[6] An alternative approach could be the assumption of double earner households in the case of partner households. The ratio would be higher for households with two adult persons. However, the simultaneous realization of two employments should be more difficult than the realization of one employment. Therefore at least for short-term utility maximization the assumption of a single earner household seems to be more realistic.

An alternative explanation for a lower exit probability for households with a lower ratio could be a lack of demand for low skilled workers. The individual market wage strongly depends on the amount of human capital. The ratio between the two income sources varies with the human capital of the household

[5]We assume the income ratio to be exogeneous. However, one could argue that demographic behavior with respect to fertility and marriage or schooling are affected by the welfare system and therefore the income ratio could be endogenous. For example, Keane and Wolpin (2002) take into account the impact of welfare benefits on the economic and demographic behavior of women in the U.S.

[6]We do not include the costs of working like child care costs in our analysis, but we control for several household characteristics in our empirical model like the existence of children.

members. Therefore an observed influence of the income ratio on the duration of welfare receipt could be caused by a relatively low arrival rate of job offers for low qualified individuals, even if we hold the job search intensity constant. Theoretically, the influence of a lower arrival rate on the exit probability is ambiguous because the reservation wage depends positively on the arrival rate. But nonetheless, a positive influence of the income ratio could also indicate a lack of demand for low skilled workers. However, different low skilled workers may have different reservation wage levels according to their household related welfare claims. Controlling for skill level may therefore allow for a discrimination between demand effects and incentive effects. In addition to that we control for the local labour market performance.

We assume that the effect of the ratio differs within the range of values. If the expected market wage is clearly below the reservation wage, the probability of an acceptable offer should be small and therefore the search behavior should differ qualitatively from households whose expected market wage is in the same range as the reservation wage or above it. If there exists a range of values corresponding with zero job search activities, an increase of the market wage in this range should not lead to a higher exit probability. Therefore we test the influence of the ratio separately for three ranges of values.

2.3 Data, Variables, and Methods

This study uses data from the German Socio-Economic Panel study (SOEP). The yearly repeated SOEP started 1984 in west Germany and was extended to include east Germany in 1990. In all panel waves, the head of the household provides information about the household and every household member aged 16 or older provides additional individual information (for details on the SOEP see Schupp and Wagner, 2002). Between 1992 and 2000, retrospective monthly information about social welfare receipt for each month of the previous calendar year is part of the household questionnaire. Excluding households with a head and if existing her partner aged 61 years or older at the beginning of the spell we

Table 2.1: Number of spells per household

number of spells	Freq.	Percent
1	357	78.46
2	76	16.70
3	19	4.18
4	2	0.44
5	1	0.22
Total	455	100

observe 579 uncensored or right-censored social welfare spells between January 1991 and December 1999, distributed on 455 households. The maximum number of spells of each household is five (one household), 357 households experience one spell of social welfare receipt (Table 2.1). These spell data are combined with several time-variant and time-invariant household and individual characteristics.

In the data there are 386 uncensored and 193 right-censored observations (Table 2.2). We are interested in the transition from social welfare to a situation with employment income. Therefore we differentiate between transitions to employment (199 cases) and alternative transitions (187 cases). A transition to employment is defined as a situation with at least one adult household member (head of the household or her partner) working full-time, both working part-time or one person working part-time in the case of single households subsequent to benefit receipt, at the latest beginning two months after the spell ending. The length of social assistance receipt in our sample ranges from 1 to 90 months. The mean spell length is 13.4 months, spells with a transition to work have an average length of 10.3 months and are on average one month shorter than spells with alternative transitions (11.6 months).

Descriptive statistics of the covariables are documented in Table 2.3. These statistics refer (a) to the status at the beginning of a welfare spell (n=579) and (b) to the monthly status (every month one observation, n=7752). Spell observations mostly ends within one year (71 %), afterwards the number of spells ending decreases constantly. About 15% of the whole sample end in the second year of observation, 7% in the third and 3% in the fourth year, 4% last for more than four years. These proportions refer to all spells, independent of the censor status.

Table 2.2: Length and Destination states, social assistance spells

Destination state	Freq.	Percent	Average Length (Standard Deviation)
Right censored	193	33.3	18.2
			(17.8)
Transitions to employment	199	34.4	10.3
			(10.3)
Alternative transitions	187	32.3	11.6
			(12.6)
Total	579	100	13.4

To control for the economic situation we include time dummies for each year of observation. The proportion of spells beginning in different years ranges from 4% in 1991 up to 16% in 1994.[7] Disproportional numbers of welfare spells start in January or end in December. Therefore we include January and December dummies in our analyses. Around one quarter of the observed households live in east Germany. The mean of the local unemployment rate is 11.6%, whereby the values range from 3.7 to 21.7 referring to federal states and yearly averages.[8] Nearly half of the households are single households, 38% female and 10% male singles. The head of a household or her partner is aged older than 50 years in 12% of the observed spells and in 32% a foreign head or partner is living in the household. In every tenth household the head or his partner is handicapped, which means that at least one of these persons answers the question whether he or she is officially registered as having a reduced capacity for work or of being severely disabled with yes. Children aged 6 years and younger live in 40% per cent of the households, children between 6 and 18 in 36%. In nearly all households the head or her partner holds at least a compulsory school degree (93%) while only in about two thirds of all households at least one of these persons has finished vocational training (60%). The statistics based on the observed months differ from the reported statistics due to the higher weight of longer spells.

[7] One has to be careful with interpretation of these descriptive statistics. For example the increase in social assistance spells beginning in 1999 can be at least partly explained by the new sub sample F ("innovation sample") of the SOEP in 2000. Due to this new sample F the sample size of the SOEP increased substantially.

[8] The unemployment rate is defined as the quotient between unemployed registered persons and persons in civilian employment. The rates are taken from the German Statistical Yearbook (Statistisches Bundesamt, 2001).

Table 2.3: Descriptive Statistics

Variable	(a) Mean/ share (standard deviation)	(b) Mean/ share (standard deviation)
End/ time of observation		
1 year	0.71	0.61
2 years	0.15	0.20
3 years	0.07	0.09
4 years	0.03	0.05
> 4 years	0.04	0.05
Year of observation		
1991	0.04	0.01
1992	0.07	0.05
1993	0.11	0.09
1994	0.16	0.12
1995	0.12	0.13
1996	0.10	0.15
1997	0.13	0.15
1998	0.12	0.16
1999	0.15	0.15
December	0.07	0.10
January	0.34	0.07
East Germany	0.27	0.21
Local unemployment rate	11.62 (4.12)	11.65 (4.09)
At least one hh-member with vocational training	0.93	0.90
At least one hh-member with school graduation	0.60	0.56
No partner (female)	0.38	0.43
No partner (male)	0.10	0.08
Household member > 50	0.12	0.20
Children aged <6	0.32	0.33
Children aged $\geqslant 6 \leqslant 18$	0.40	0.43
Non German hh-member	0.36	0.34
Handicapped hh-member	0.10	0.13
Income Ratio >1	1.39 (0.50)	1.33 (0.44)
Number of observations	579	7,752

Source: SOEP, numbers refer to (a) first month of each spell, and (b) to all observed months, standard deviations in parentheses.

Before discussing the ratio and the difference between the potential household income in case of one adult person working full-time and the social assistance amount, we describe the estimating and calculating procedures of these two income sources separately in the following.

2.3.1 Estimation of Potential Net-Income

In a first step we estimate potential gross market wages of all heads of the household and as the case may be of their partner. We cannot observe their wages directly because most of the individuals in our data set are not working while receiving social assistance. Therefore we estimate the potential wages using all individuals in working age. Whether or not we observe wages depends on an individual's participation decision. Due to this self-selection we cannot assume the sample of workers to be a random sample of all potential working individuals and we have to account for the sample selection problem.

The sample selection model we apply, also referred to as the type II Tobit model (e.g. Wooldridge, 2002, or Greene, 2003), consists of a log-linear wage equation

$$\ln w_i = X_{1i}\beta_1 + \epsilon_{1i} \tag{2.1}$$

with X_{1i} as a vector containing exogenous characteristics and w_i as person's i wage and an equation describing the binary choice to work or not to work and therefore determining the sample selection

$$z_i^* = X_{2i}\beta_2 + \epsilon_{2i} \tag{2.2}$$

We observe wages according to the rule:

$$\begin{aligned} w_i = w_i^*, \quad z_i = 1 \text{ if } z_i^* > 0 \\ w_i \text{ not observed, } z_i = 0 \text{ if } z_i^* \leq 0 \end{aligned}$$

whereby z_i indicates working or not working and this depends on the characteristics X_{2i}. One can estimate the wage equation consistently assuming that the two error components of the two equations follow a bivariate normal distribution. The expected value of ln(wage) for individuals not working corresponds to:

$$E\{\ln w_i | z_i = 0\} = X_{1i}\beta_1 - \sigma_{12}\frac{\phi(X_{2i}\beta_2)}{1 - \phi(X_{2i}\beta_2)} \tag{2.3}$$

We estimate separated models for east and west Germany and for men and women with a pooled sample using the SOEP waves from 1991 - 1999.[9] The estimation results are reported in Tables 2.6 and 2.7. We control for the year and the region. Education, measured in years, age and firm specific capital, measured in years being employed at the actual employer, have significantly positive influence on the wage per hour, while the squared age and the squared firm specific capital have a significantly negative impact. Foreigners have lower wages in both regions and the absence from the labor market in years, accounting for the previous five years, have a negative impact on the wage. While the squared absence from the labor market influences the wage positively in east Germany, the effect in west Germany is insignificant.

Using these estimation results, we calculate a potential monthly full-time gross wage for each head of household and her partner. Calculating the potential net income, we assume that in the case of a partner household the person with the higher income would work and we account for income taxes, social security contributions, child and housing allowance.

2.3.2 Social Assistance

The amount of social assistance was not asked in all waves of the SOEP. Furthermore, in the years the amount of social assistance was part of the questionnaire, the current amount but not the monthly amount during the previous year was asked. Therefore we can observe the monthly receipt as a binary variable but not the corresponding amount of social assistance.

Instead of direct observation we calculate the maximum of social assistance.

[9]We excluded individuals working short time (*Kurzarbeit*), doing a vocational training, military or community service. In addition to that we exclude persons aged 60 and older.

This is the permanent income, including other transfers, a household staying on welfare would receive permanently. As described above this amount depends on the number and the age of household members and varies by the region and the year of receipt. We use the average yearly individual basic allowances for east and west Germany to calculate the basic allowance for each household member and add them up. Moreover we consider the one-time payments by using the same method as (Breuer and Engels, 2003) or (Boss, 2002): We calculate 16% of the individual basic allowance for the head of household, 17% for the partner and 20% for each child. In addition to that we take an allowance for housing depending on the household size into account.

2.3.3 Ratio between Employment Income and Social Assistance

We calculate the ratio between the potential household net income in case of one person working fulltime and the amount of transfer payment. The empirical distribution of this variable in the first month of each spell is plotted in Figure 2.2. The median is 1.26, i.e. for about half the sample expected income does not exceed their welfare benefits by more than 25%. This indicates that the incentives to search for a job may be low for a lot of individuals being on social welfare. The median of the distribution corresponding to all observed months is lower with 1.21 (see appendix, Figure 2.6), which reflects the higher weight of longer spells in the distributions of all month-observations. This indicates that a lower income ratio may go along with a longer stay in the social assistance.[10] We interact the ratio with three dummy variables and thereby split the ratio in three parts:

- ratio1: takes on the ration value if the ratio is below 1 (25% of the sample)
- ratio2: takes on the ration value if the ratio is between 1 and 1.5 (45%)
- ratio3: takes on the ration value if the ratio is above 1.5 (29%)

[10]Separated histograms for single and couple households are presented in the Appendix, Figures 2.7 and 2.8.

One could argue that the difference between potential household net income could never be negative and therefore the ratio could never be lower one, because these households would receive supplementary transfer payments (see for example Wilde, 2003). Nevertheless we use ratios lower than one in our analysis, because we estimate the mean of a wage distribution an individual is faced with and not a deterministic wage. It is possible that a person receives a job offer with a wage resulting in a higher net household income than social assistance, although the mean of his wage distribution is lower than the benefits. A ratio lower than one indicates a relatively low probability of such an offer. Censoring the ratio variable would lead to a loss of information indicating the probability of acceptable offers.

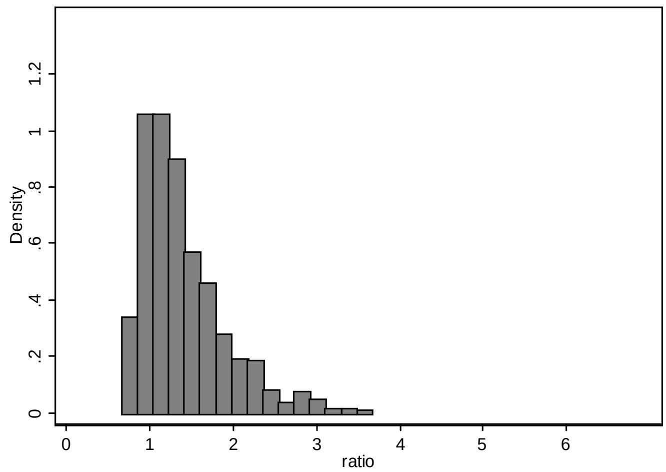
In addition to that one can observe households who are eligible for social assistance but do non take it up. This (non-) take-up behavior depends among others on the expected benefit amount (see e.g. Riphahn, 2001): The probability of take-up rises with the potential amount of transfer payments. Because we are interested in the leaving processes of social assistance, the ratio between the two separate income sources and not the combination of the different income sources is the relevant variable.

In Figure 2.3 we plot the corresponding income ratio of the single earner households not receiving social welfare in the SOEP. The median of this distribution is with 1.89 clearly higher than the one of the estimated ratio of households receiving social welfare. 2.7 % of these households have a ratio lower than one, i.e. they would have a higher income receiving social welfare instead of one person working full time. On the one hand this could be explained with application and stigma costs (see e.g. Kayser and Frick, 2001, Riphahn, 2001, or Wilde and Kubis, 2005). On the other hand we only account for employment income and do not consider other income. Therefore the net household income used for the calculation is the lower bound of the real net income.

2.3.4 Model Specification

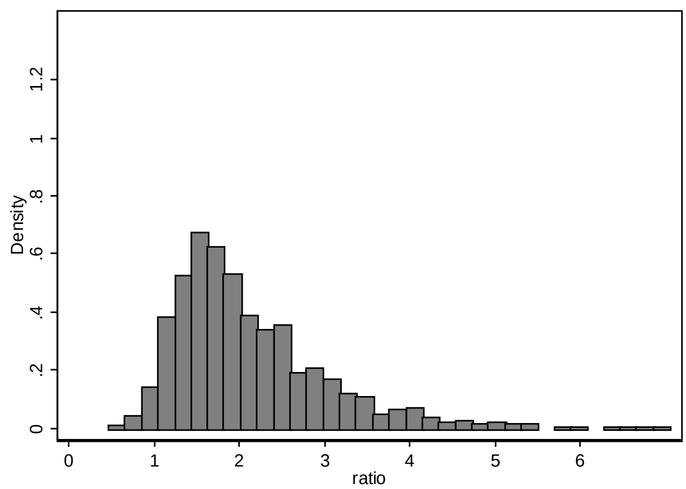
The process of leaving social welfare in favor of labor income can appropriately be modeled by a transition rate approach. According to the type of data being used

Figure 2.2: Histogram: Ratio between potential net income and the amount of social assistance, first month of welfare spell



Source: SOEP, waves 1991-1999, n=579

Figure 2.3: Histogram: Ratio between potential neet income and the amount of social assistance for single earner households in the SOEP in 1999



Source: SOEP, wave 1999, n=2411

here, a discrete hazard rate model has to be applied (see for example Han and Hausman, 1990, Jenkins, 2004, Meyer, 1990, Sueyoshi, 1992, Narendranathan and Stewart, 1993a). The duration of welfare receipt is generated by a continuous time process, but observed or grouped in monthly intervals. Two potential destination states q are considered reflecting transitions to employment ($d = 1$) and alternative transitions like for example other transfer programs or marriage ($d = 2$). The overall hazard rate is defined as the limit of the conditional probability for the ending of a spell in interval $[t; t+t[$ given that no transition occurred before the start of this interval:

$$\lambda_s(t) = \lim_{\triangle t \to 0} \frac{P(t \leq T_s \leq t + \triangle t \mid T_s \geq t)}{\triangle t} \tag{2.4}$$

where T denotes the length of a spell. T is assumed to be a continuous, non-negative random variable. We assume proportional transition rates with covariates causing proportional shifts of a so-called baseline transition rate and interval constant covariates. The hazard rate $\lambda_{sd}(t)$ corresponds to the sum of the two transition rates[11]

$$\lambda(t \mid x_i(t), \eta_i) = \sum_{d=1}^{2} \lambda_d(t \mid x_i(t), \eta_{id})$$

with the transition rate to destination state d corresponding to

$$\begin{aligned} \lambda_d(t \mid x_i(t), \eta_{id}) &= \lambda_{0d}(t) \exp(x_i(t)\beta_d + \eta_{id}); \\ (\eta_1, \eta_2) &\sim N(0, 0, \sigma_1^2, \sigma_2^2, \rho) \end{aligned} \tag{2.5}$$

$\lambda_{0d}(t)$ denotes the destination specific baseline transition rate, $x_i(t)$ an individual time variant row vector of covariates for individual i, β_d a column vector of parameters and η_{id} a time invariant individual and destination specific error term, representing the joint influence of unobserved heterogeneity. We assume these error terms or random intercepts to independent of the observed individual

[11] In principle destination specific covariables x_q are allowed but not considered here. For the general model with destination specific covariables see appendix.

characteristics and to be bivariate normally distributed with expected values 0, which allows for dependent competing risks.

We observe the duration of unemployment and employment in monthly intervals. This implies that instead of continuous levels of $x_i(t)$ their interval specific levels have to be taken into account. Assumed that the time axis is divided into intervals of unit length, a given spell consists of a number of j intervals, in the following referred to as subspells. The interval specific levels of $x_i(t)$ and the observed interval baseline hazard $\lambda_{0d}(t)$ for the $k-th$ subspell are denoted as x_{ik} and $h_{0d}(k)$.

For the survivor function in social assistance this implies:

$$S_(j|x_i, \eta_i) = \exp\Big(-\sum_{d=1}^{2}\sum_{k=1}^{j}\exp(x_{ik}\beta_d + h_{0d}(k) + \eta_{id})\Big)$$
$$h_{0d}(k) = \ln\Big(\int_{t_{k-1}}^{t_k}\lambda_{0d}(\tau)d\tau\Big). \tag{2.6}$$

The survivor function $S(j)$ describes the probability that a spell lasts at least j intervals. The h_0 parameters are capturing the duration dependence of the baseline transition function. They may be interpreted as an interval specific mean of the baseline transition rate, which is equivalent to an interval specific constant baseline transition rate.

Following from this, the probability h of a transition to state r at a given interval j is given by the difference of two survivor functions multiplied by the share of the risk-specific transition rate at interval j related to the hazard rate at interval j.

$$h_r(j) = \frac{\exp(x_{rj}\beta_r + h_{0r}(j) + \eta_r)}{\sum_{d=1}^{2}\exp(x_{dj}\beta_d + h_{0d}(j) + \eta_d)}[S(j-1) - S(j)] \tag{2.7}$$

The likelihood contribution of a spell corresponds to

$$l(\beta, h_0, \eta) = \frac{\exp(x_{1j}\beta_1 + h_{01}(j) + \eta_1)^{c_1} \exp(x_{2j}\beta_2 + h_{02}(j) + \eta_2)^{c_2}}{\sum_{d=1}^{2} \exp(x_{dj}\beta_d + h_{0d}(j) + \eta_d)} [cS(j-1) - (2c-1)S(j)] \quad (2.8)$$

whereby $c_1 = 1$ and $c_2 = 1$ indicate a transition to risk 1 and risk 2 in interval j, respectively, and c corresponds to the maximum of $c1$ and $c2$. It implies that right-censored spells are assumed to be censored at the end of the related interval, but that transitions may occur somewhere between $j-1$ and j. The likelihood contribution is not separable into destination-specific components as suggested by Narendranathan and Stewart (1993b) because we do not assume that transitions can only occur at the interval boundaries (see Roed and Nordberg, 2003, or Jenkins, 2004, for similar approaches. Therefore we can not estimate destination specific models separately, even in a model without unobserved heterogeneity. In the following, we will refer to this as a piecewise exponential model and a random effects piecewise exponential model, respectively. The likelihood function is solved by applying Gauss-Hermite quadrature.

2.4 Results

We estimate discrete time hazard rate models with and without unobserved heterogeneity. The coefficients can be interpreted with respect to the underlying continuous time proportional hazard rate. We estimate our models with and without splitting the income ratio into three parts, the results are reported in Table 2.4 and Table 2.5 respectively. The inclusion of unobserved heterogeneity does not significantly improve the model fit.

We created three variables representing the effect of the ratio between estimated potential labor income and welfare payment level: The first for ratios lower and equal 1, the second for ratios from 1 to 1.5 and the third for ratios above 1.5.[12] In both models with and without unobserved heterogeneity the coefficients

[12]The ratio takes on the value 1 if the potential labor income equals the welfare payment level and the value 1.5 if the potential labor exceeds the social welfare payments by 50%.

of the latter two variables are significantly positive, while the coefficient of the first is positive but not significantly different from zero. An increase in the ratio seems to be more relevant if the potential labor income exceeds the social assistance level (see Table 2.4). Estimating the model with one ratio variable leads to a weakly significantly positive influence (at the 10 percent level, see Table 2.5).

The coefficient of the income ratio for a ratio between 1 and 1.5 indicates that a 0.1 higher ratio goes along with a 10% higher probability of an exit to employment, while a 0.1 higher ratio for ratios above 1.5 leads to 7% higher probability of a transition to employment. However, the difference between the two coefficients is not significant and therefore a further interpretation is not useful. Assuming households with the same welfare level, a difference in the income ratio of 0.1 stands for a difference in estimated labor income by 10% of the social welfare level. For alternative transitions, these income variables have no significant influence. Our results confirm our predictions: Given a level of social welfare payment it is more likely to observe exits from social assistance to employment, the higher an individuals (net) market wage is. This is especially the case for households with an expected labor income higher than the social assistance level. Only if the household is able to improve its income through employment, the difference between the two income sources matter. The hazard rates of two types of households with income ratios of 0.5 and 2 are plotted exemplarily in Figure 2.9 in the Appendix. The hazard rates are calculated for average households, differing only in their income ratios. The estimated hazard rate of the household with an income ratio of 2 is 2.3 times higher than the one of the household with the lower income ratio.

The relevance of incentive effects is stressed by the fact that skill indicators and the local unemployment rate turn out to be insignificant.[13] The other relevant covariates for the transition from welfare to work are quite similar, independent of the model we estimate. Households with a head being single have a significant lower probability of leaving social welfare via employment than partner house-

[13]One could argue that the local unemployment rate is endogenous because the transitions from social assistance to work directly influence the rate of unemployed persons. However, the results do not change leaving out the unemployment rate.

holds. This effect is especially strong for women. Households with the head or her partner being older than 50 years have a lower exit probability than younger households. The presence of young children has no significant effect on the duration of welfare receipt, while older children between 6 and 18 reduce the duration of social welfare receipt. Households in east Germany exit faster to employment, which is a surprising result because of the relatively bad economic performance of east Germany. One possible (ad-hoc) explanation may be a relatively large number of transitions into public financed jobs for unemployed persons in east Germany (e.g. *Arbeitsbeschaffungsmaßnahmen* or *Strukturanpassungsmaßnahmen*), but this has to be checked empirically. The existence of a handicapped adult household member seems to have no influence on the transition probability. Moreover the nationality of adult household members does not affect the exit probability of households. In addition to that the existence of an adult person with vocational or a school graduation has no influence on the probability of exiting social welfare. This result is similar to that of Riphahn (1999) who identifies only a significant effect for a university degree but not for vocational training while Wilde (2003) and Gangl (1998) identify positive effects of a vocational training.[14] The local unemployment rate has a negative but insignificant effect on the welfare duration. We re-estimated the model with different discrete time duration models resulting from different assumptions about the underlying continues time process (multinomial logit models and complementary log-log models). The results do not change qualitatively.[15] This is line with the results of a monte carlo study by Jenkins (2004), indicating that a "wrong" specification of the discrete time duration models leads to a relevant bias only if the discrete intervals are relatively wide.

In contrast to the transitions to work, the income variables have no sig-

[14]Riphahn estimates two sorts of models using different covariables: Duration models with continuous time and household as well as individual characteristics on the one hand and duration models with discrete time and household characteristics with an additional estimated income variable on the other hand. We refer to both model categories.

[15]Additionally the we applied simulation techniques instead of Gauss-Hermite to solve the two-dimensional integral. However, the models did not converge, which can be explained by the fact that there seems to be no unobserved heterogeneity in the data. Therefore, the non-convergence of these models is in line with the results presented here.

nificant influence on the probability of alternative transitions, reported in the rows "Alternative Transitions" in the Tables 2.4 and 2.5. This is an expected result and shows the importance to differentiate between alternative risks when examining the transition from welfare to work and the role of estimated labor income.

Our results confirm our predictions: Given a level of social welfare payment it is more likely to observe exits from social assistance to work, the higher an individuals (net) market wage is. In contrast to other studies like Riphahn (1999) or Wilde (2003) we estimate a positive effect of the potential net labor income on the transition probability and this effect seems to be more relevant for households with a potential market wage above their social assistance level.

In the SOEP the net household income is asked every year. In Figure 2.4 both, the predicted and the realized income ratios of the households leaving social welfare for work are plotted.[16] The mean of the realized ratio distribution is with 1.43 slightly above the estimated ratio with a mean of 1.36. Around 20% of our households have a lower income after leaving social welfare. This observation is in line with the results of Wilde (2003) who observes 25% of the households realizing a lower income. There exist several possible explanations for this observation. Because of stigma costs it could be rational for some households to accept an income loss, there could exist measurement errors in the income variable or some households could earn additional money with unobservable illegal employment. However, most of our observation have a significantly higher income after leaving social welfare for work. In Figure 2.5 our calculated and the realized ratios are plotted together. One can see that both variables correlate not perfectly but clearly positively.

[16] In 49 cases the income variable is missing or the household started to receive social assistance again.

Table 2.4: Discrete-time proportional hazard rate models, three ratio variables

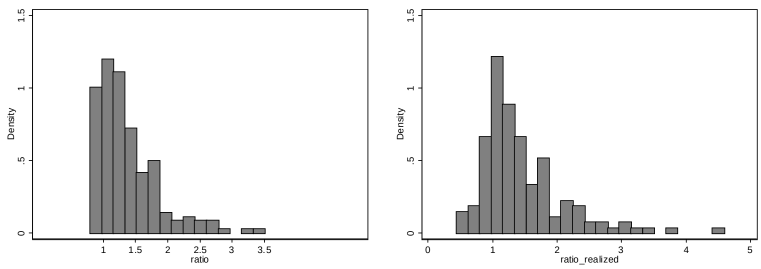
	Piecewise exponential model				Random effects piecewise exponential model			
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Transitions to employment		Alternative transitions		Transitions to employment		Alternative transitions	
2 years	-0.41	0.22	-0.57*	0.22	-0.37	0.28	-0.41	0.36
3 years	-0.31	0.32	-0.66	0.34	-0.27	0.42	-0.40	0.55
4 years	-0.79	0.52	-0.75	0.47	-0.74	0.62	-0.45	0.70
5 and more years	-1.30	0.73	-0.97	0.60	-1.25	0.83	-0.63	0.84
Year of observation								
1992	-0.41	0.53	0.25	0.54	-0.42	0.54	0.21	0.57
1993	-0.12	0.48	0.31	0.52	-0.13	0.50	0.26	0.56
1994	0.02	0.46	0.41	0.52	0.03	0.49	0.42	0.55
1995	-0.29	0.47	0.10	0.52	-0.28	0.49	0.10	0.55
1996	-0.48	0.49	0.41	0.53	-0.48	0.52	0.4	0.57
1997	-0.25	0.47	0.28	0.54	-0.25	0.52	0.23	0.59
1998	0.14	0.46	0.28	0.54	0.14	0.51	0.27	0.58
1999	-1.22*	0.51	-0.75	0.57	-1.22*	0.55	-0.81	0.61
December	2.32**	0.14	3.03**	0.16	2.33**	0.15	3.08**	0.19
January	-1.44*	0.71	-0.55	0.59	-1.44*	0.71	-0.52	0.59
East Germany	0.77*	0.22	0.73*	0.36	0.78*	0.34	0.79	0.40
Local unemployment rate	-0.05	0.04	-0.05	0.04	-0.05	0.04	-0.05	0.04
At least one hh-member with vocational training	0.20	0.19	-0.17	0.18	0.20	0.19	-0.17	-0.85
At least one hh-member with school graduation	0.16	0.37	0.26	0.31	0.17	0.38	0.29	0.84
No partner (female)	-0.68**	0.18	0.15	0.18	-0.68**	0.19	0.18	0.80
No partner (male)	-0.60	0.35	0.13	0.31	-0.61	0.36	0.15	0.40
Household member > 50	-0.83**	0.29	-0.42	0.26	-0.84**	0.30	-0.49	-1.60
Children aged <6	-0.26	0.17	-0.14	0.19	-0.27	0.17	-0.15	0.20
Children aged $\geqslant 6 \leqslant 18$	0.46*	0.18	0.08	0.20	0.46*	0.19	0.08	0.22
Non German hh-member	-0.23	0.18	0.17	0.18	-0.23	0.19	0.24	0.25
Handicapped hh-member	-0.15	0.27	-0.02	0.26	-0.16	0.28	-0.04	0.29
Income Ratio >1	0.92	0.57	0.56	0.56	0.91	0.58	0.50	0.62
$1 \leqslant$ Income Ratio < 1.5	0.93*	0.42	0.33	0.42	0.93*	0.42	0.30	0.46
$1.5 \leqslant$ Income Ratio	0.64*	0.28	0.41	0.27	0.64*	0.28	0.42	0.31
Constant	-4.28**	0.80	-4.94	0.81	-4.29**	0.81	-5.11**	1.00
$\text{Ln}(\sigma^2)$	-	-	-	-	-4.33	57.81	-0.49	1.29
$\text{cov}(\eta_1, \eta_2$		-	-			-0.18	0.67	
Log-Likelihood	-1,414.42				-1,414.24			

579 spells, 7752 months, the unobserved heterogeneity is assumed to follow a bivariate normal distribution. *: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level. For missing values concerning the handicap and vocational training variables additional dummies are included. Their insignificant coefficients are not reported here.

2.5 Conclusion

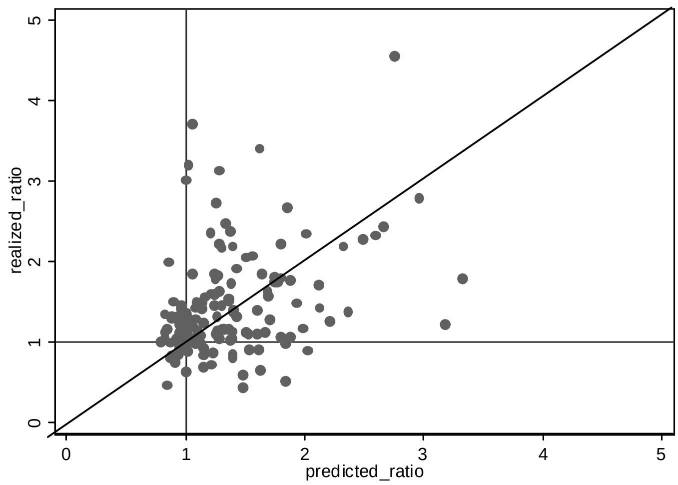
The aim of the study was to estimate the influence of the ratio between estimated potential labor income and the welfare payment level on the probability of a transition from social welfare to work. We use data from the SOEP waves 1992-2000

Figure 2.4: Histogram: Ratio between the net income and the amount of social assistance: predicted (left) and realized (right)



Source: SOEP, waves 1991-2000, n=199 (predicted) and n=150 (observed). In 49 cases the income variable is missing or the household started to receive social assistance again.

Figure 2.5: Scatterplot of the predicted and the realized ratio



Source: SOEP, waves 1991-2000, n=150.

including information about spell duration of households receiving social welfare and the monthly employment status of the household members. The potential net labor income is estimated with standard wage equations accounting for sample selection and applying a simple tax function. We estimate a discrete-time proportional hazard rate model with competing risks and risk specific unobserved heterogeneity.

The ratio between potential labor income and the welfare level shows a positive effect on the probability of a transition to employment for households whose potential labor income exceeds their welfare payment level. Our results are contrary to previous studies dealing with the determinants of welfare spell duration in Germany: We identify an effect of the income ratio according to the standard theoretical predictions. This "new" result derives from a simultaneous consideration of both sources of income, the net household labor income and the social welfare level, and additionally from a differentiation between transitions to work and alternative transitions.

The alternative explanation for low skilled workers being more likely to be hit by long term unemployment according to a lower job offer arrival rate for low skilled employment turns out to be of minor relevance. Contrary to the ratio indicators, skill indicators are far from being significant. Obviously, the explanatory power of skills is outweighed by incentive effects.

A reduction of the benefit level could be one solution to overcome the incentive problems. However, the amount of social assistance is related to a basic minimum income concept and a general reduction of the benefit level seems to be no political option. But a reduction of the social assistance level is not the only way to overcome incentive problems of a transfer program, there exist other possible solutions like for example workfare.

2.6 Appendix

Random effects piecewise exponential model

Assumption: Proportional hazard rate model with two competing risks and unobserved heterogeneity

$$\lambda(t \mid x, \eta) = \lambda_1(t \mid x_1, \eta) + \lambda_2(t \mid x_2, \eta)$$
$$\lambda_1(t \mid x, \eta) = \lambda_{01}(t) \exp(x_1(t)\beta_1 + \eta_1); \quad \lambda_2(t \mid x, \eta) = \lambda_{02}(t) \exp(x_2(t)\beta_2 + \eta_2)$$

With the assumption of interval constant covariates x_d it follows:

$$\begin{aligned}
S_(j|x, \eta) &= \exp\Big(- \int_0^{t_j} \lambda(\tau) d\tau \Big) = \exp\Big(- \sum_{k=1}^{j} \int_{t_{k-1}}^{t_k} \lambda(\tau) d\tau \Big) \\
&= \exp\Big(- \sum_{d=1}^{2} \sum_{k=1}^{j} \int_{t_{k-1}}^{t_k} \lambda_{0d}(\tau) \exp(x_{dk}\beta_d + h_{0d}(k) + \eta_d) d\tau \Big) \\
&= \exp\Big(- \sum_{d=1}^{2} \sum_{k=1}^{j} \exp(x_{dk}\beta_d + h_{0d}(k) + \eta_d) \Big); \quad h_{0d}(k) = \ln\Big(\int_{t_{k-1}}^{t_k} \lambda_{0d}(\tau) d\tau \Big)
\end{aligned}$$

In the following $\gamma_{dk} = exp(x_{dk}\beta_d + h_{0d}(k) + \eta_d)$. Assuming interval constant transition rates $\lambda_{0d}(k)$ the transition probability for a destination state $r = 1, 2$ corresponds to:

$$\begin{aligned}
h_r(j) &= \gamma_r \int_{t_{j-1}}^{t_j} \exp\Big(- \sum_{d=1}^{2} \Big(\Big(\sum_{k=1}^{j-1} \gamma_d \Big) + \int_{t_{j-1}}^{\tau} \lambda_{0d}(u) \exp(x_{dj}\beta_d + \eta_d) du \Big) \Big) d\tau \\
&= \gamma_r \exp\Big(- \sum_{d=1}^{2} \sum_{k=1}^{j-1} \gamma_d \Big) \int_{t_{j-1}}^{t_j} \exp\Big(- \sum_{d=1}^{2} \gamma_d \int_{t_{j-1}}^{\tau} du \Big) d\tau \\
&= \gamma_r \exp\Big(- \sum_{d=1}^{2} \sum_{k=1}^{j-1} \gamma_d \Big) \int_{t_{j-1}}^{t_j} \exp\Big(- \sum_{d=1}^{2} (\tau - t_{j-1}) \gamma_d \Big) d\tau \\
&= \gamma_r \exp\Big(- \sum_{d=1}^{2} \sum_{k=1}^{j-1} \gamma_d \Big) \Big[\frac{- \exp\Big(- \sum_{d=1}^{2} (\tau - t_{j-1}) \gamma_d \Big)}{\sum_{d=1}^{2} \gamma_d} \Big]_{t_{j-1}}^{t_j}
\end{aligned}$$

$$
\begin{aligned}
h_r(j) \quad &= \frac{\gamma_r}{\sum_{d=1}^{2}\gamma_d}\exp\Big(-\sum_{d=1}^{2}\sum_{k=1}^{j-1}\gamma_d\Big)\Big[1-\exp\Big(-\sum_{d=1}^{2}\gamma_d\Big)\Big] \\
&= \frac{\gamma_r}{\sum_{d=1}^{2}\gamma_d}\Big[\exp\Big(-\sum_{d=1}^{2}\sum_{k=1}^{j-1}\gamma_d\Big)-\exp\Big(-\sum_{d=1}^{2}\gamma_d\Big)\Big] \\
&= \frac{\gamma_r}{\sum_{d=1}^{2}\gamma_d}[S(j-1)-S(j)]
\end{aligned}
$$

This leads to a likelihood function which is not separable into destination-specific components because we do not assume that transitions can only occur at the interval boundaries. Therefore we can not estimate destination specific models separately, even in a model without unobserved heterogeneity.

The likelihood contribution of a single observation can be written as:

$$
\begin{aligned}
l(\beta, h_0, \eta) = \quad & \frac{\exp(x_{1j}\beta_1 + h_{01}(j) + \eta_1)^{c_1}\exp(x_{2j}\beta_2 + h_{02}(j) + \eta_2)^{c_2}}{\sum_{d=1}^{2}\exp(x_{dj}\beta_d + h_{0d}(j) + \eta_d)} \\
& [cS(j-1) - (2c-1)S(j)]
\end{aligned}
$$

$$
\begin{aligned}
\text{with} \qquad & c_1 = \begin{cases} 1, \text{in the case of a transition to the destination state 1 in interval j} \\ 0, \text{otherwise} \end{cases} \\
& c_2 = \begin{cases} 1, \text{in the case of a transition to the destination state 1 in interval j} \\ 0, \text{otherwise} \end{cases} \\
& c = \max(c_1, c_2)
\end{aligned}
$$

The overall likelihood function is:

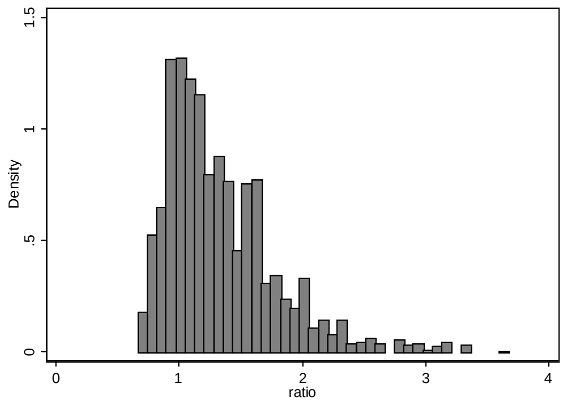
$$
\begin{aligned}
l(\beta, h_0, \eta) = \quad & \prod_{i=1}^{n}\frac{\exp(x_{1j}\beta_1 + h_{01}(j) + \eta_1)^{c_1}\exp(x_{2j}\beta_2 + h_{02}(j) + \eta_2)^{c_2}}{\sum_{d=1}^{2}\exp(x_{dj}\beta_d + h_{0d}(j) + \eta_d)} \\
& [cS(j-1) - (2c-1)S(j)]
\end{aligned}
$$

Table 2.5: Discrete-time proportional hazard models, one ratio variable

	Piecewise exponential model				Random effects piecewise exponential model			
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Transitions to employment		Alternative transitions		Transitions to employment		Alternative transitions	
2 years	-0.41	0.21	-0.56*	0.22	-0.36	0.27	-0.35	0.34
3 years	-0.35	0.32	-0.63	0.34	-0.27	0.40	-0.31	0.52
4 years	-0.84	0.52	-0.72	0.47	-0.75	0.59	-0.34	0.66
5 and more years	-1.36	0.73	-0.92	0.60	-1.25	0.80	-0.49	0.81
Year of observation								
1992	-0.40	0.53	0.21	0.54	-0.42	0.55	0.17	0.58
1993	-0.10	0.49	0.26	0.52	-0.10	0.50	0.21	0.56
1994	0.06	0.48	0.37	0.51	0.07	0.49	0.38	0.56
1995	-0.27	0.49	0.06	0.52	-0.26	0.50	0.07	0.57
1996	-0.44	0.51	0.34	0.52	-0.43	0.52	0.34	0.57
1997	-0.18	0.52	0.20	0.54	-0.18	0.53	0.14	0.59
1998	0.19	0.50	0.19	0.53	0.20	0.51	0.19	0.58
1999	-1.16*	0.54	-0.84	0.56	-1.17*	0.55	-0.90	0.61
December	2.32**	0.14	3.03**	0.16	2.34**	0.15	3.10**	0.19
January	-1.44*	0.71	-0.54	0.59	-1.44*	0.71	-0.51	0.59
East Germany	0.75*	0.33	0.76*	0.36	0.77*	0.35	0.82*	0.41
Local unemployment rate	-0.05	0.03	-0.05	0.04	-0.05	0.04	-0.04	0.04
At least one hh-member with vocational training	0.20	0.19	-0.19	0.18	0.20	0.19	-0.19	0.21
At least one hh-member with school graduation	0.17	0.37	0.23	0.31	0.18	0.38	0.28	0.36
No partner (female)	-0.69**	0.18	0.17	0.18	-0.72**	0.20	0.21	0.21
No partner (male)	-0.66*	0.34	0.15	0.31	-0.69	0.37	0.19	0.37
Household member > 50	-0.83**	0.29	-0.42	0.26	-0.85**	0.30	-0.50	0.31
Children aged <6	-0.24	0.17	-0.15	0.18	-0.25	0.18	-0.16	0.20
Children aged $\geqslant 6 \leqslant 18$	0.47*	0.18	0.11	0.19	0.48*	0.19	0.10	0.22
Non German hh-member	-0.25	0.19	0.18	0.18	-0.26	0.19	0.28	0.24
Handicapped hh-member	-0.17	0.27	-0.01	0.26	-0.19	0.28	-0.03	0.30
Income Ratio	0.45	0.23	0.40	0.23	0.45	0.24	0.45	0.27
Constant	-3.79**	0.68	-4.91**	0.70	-3.84	0.93	-5.28	0.71
$\mathrm{Ln}(\sigma^2)$	-	-	-	-	-1.08	57.81	-0.28	1.29
$\mathrm{cov}(\eta_1, \eta_2$		-	-			-0.001	0.67	
Log-Likelihood	-1,416.31				-1,415.93			

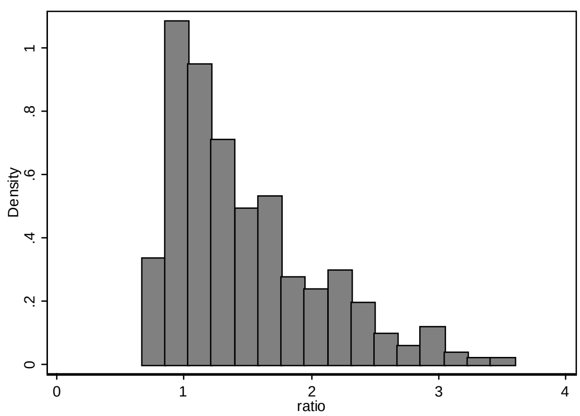
579 spells, 7752 months, the unobserved heterogeneity is assumed to follow a bivariate normal distribution. *: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level. For missing values concerning the handicap and vocational training variables additional dummies are included. Their insignificant coefficients are not reported here.

Figure 2.6: Histogram: ratio between potential net income and the amount of social assistance, all months



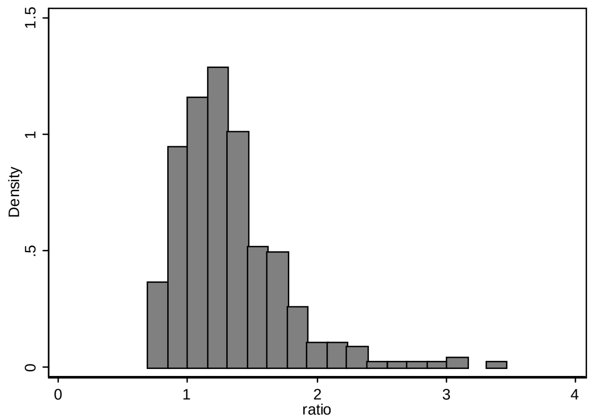
Source: SOEP, waves 1991-1999, n=7752

Figure 2.7: Histogram: ratio between potential net income and the amount of social assistance, fist month of welfare spell, single households



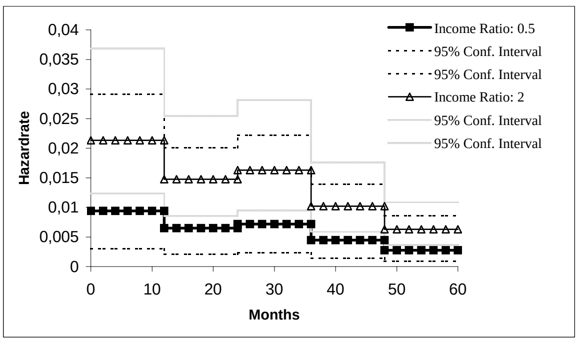
Source: SOEP, waves 1991-1999, n=277

Figure 2.8: Histogram: ratio between potential net income and the amount of social assistance, first month of welfare spell, partner households



Source: SOEP, waves 1991-1999, n=302

Figure 2.9: Hazard rates of two households with income ratio 0.5 and 2



The hazard rates are calculated for average households, differing only in their income ratios.

Table 2.6: Type II Tobit model: selection equation

	East Germany				West Germany			
	Women		Men		Women		Men	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Education in years	0.062	0.005	0.013	0.005	0.042	0.003	-0.003	0.003
Age	0.245	0.005	0.206	0.006	0.195	0.004	0.199	0.004
Age squared	-0.003	0.000	-0.003	0.000	-0.003	0.000	-0.003	0.000
Foreigner	-0.090	0.064	-0.415	0.086	-0.040	0.018	-0.060	0.019
Children < 6	0.500	0.020	-0.012	0.022	-0.629	0.015	-0.002	0.014
Partner living in hh	0.068	0.023	0.170	0.028	-0.311	0.017	0.262	0.018
Sachsen	-0.150	0.037	0.050	0.043	-	-	-	-
Sachsen-Anhalt	-0.153	0.040	0.072	0.045	-	-	-	-
Thüringen	-0.171	0.040	0.067	0.046	-	-	-	-
Mecklenburg-Vorpommern	-0.258	0.043	-0.052	0.049	-	-	-	-
Brandenburg	-0.236	0.040	-0.082	0.046	-	-	-	-
Schleswig-Holstein	-	-	-	-	-0.001	0.050	-0.032	0.051
Hamburg	-	-	-	-	0.121	0.062	0.158	0.066
Niedersachsen	-	-	-	-	-0.089	0.040	-0.095	0.041
Bremen	-	-	-	-	-0.295	0.079	0.112	0.073
Nordrhein-Westfalen	-	-	-	-	-0.141	0.037	-0.089	0.040
Hessen	-	-	-	-	0.040	0.041	0.047	0.042
Rheinland-Pfalz	-	-	-	-	-0.110	0.042	-0.045	0.044
Baden-Würtemberg	-	-	-	-	0.004	0.038	0.057	0.040
Bayern	-	-	-	-	-0.007	0.038	0.019	0.040
1992	-0.169	0.035	-0.143	0.041	0.039	0.029	-0.021	0.032
1993	-0.248	0.036	-0.267	0.042	0.023	0.029	0.005	0.032
1994	-0.263	0.038	-0.220	0.042	0.024	0.029	-0.019	0.031
1994	-0.245	0.038	-0.174	0.042	-0.000	0.029	0.004	0.032
1994	-0.257	0.037	-0.246	0.043	0.014	0.027	-0.013	0.030
1994	-0.282	0.038	-0.289	0.043	-0.032	0.028	-0.055	0.030
1994	-0.308	0.038	-0.331	0.044	-0.008	0.028	-0.056	0.030
1994	-0.274	0.038	-0.252	0.044	0.028	0.027	-0.015	0.030
Constant	-4.259	0.115	-3.300	0.126	-3.421	0.079	-3.176	0.087
Number of observations	13851		13,049		30,212		29,120	
Censored observations	6,562		5,335		16,715		11,004	

Table 2.7: Type II Tobit model: wage equation

	East Germany				West Germany			
	Women		Men		Women		Men	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Education in years	0.078	0.002	0.059	0.001	0.069	0.002	0.064	0.001
Age	0.061	0.008	0.013	0.003	0.045	0.003	0.036	0.003
Age squared	-0.001	0.000	-0.000	0.000	-0.001	0.000	-0.000	0.000
Absence from the labor market	-0.082	0.016	-0.155	0.019	-0.048	0.009	-0.076	0.015
Absence from the lab. m. squared	0.001	0.005	0.008	0.010	-0.001	0.003	-0.001	0.007
Firm-specific human capital	0.016	0.002	0.000	0.001	0.018	0.001	0.010	0.001
Firm-specific hc squared	-0.000	0.000	0.000	0.000	-0.000	0.000	-0.000	0.000
Foreigner	-0.076	0.025	0.039	0.037	-0.065	0.009	-0.040	0.007
Sachsen	-0.206	0.017	-0.195	0.016	-	-	-	-
Sachsen-Anhalt	-0.171	0.017	-0.164	0.017	-	-	-	-
Thüringen	-0.200	0.018	-0.199	0.017	-	-	-	-
Mecklenburg-Vorpommern	-0.127	0.020	-0.132	0.018	-	-	-	-
Brandenburg	-0.157	0.019	-0.141	0.017	-	-	-	-
Schleswig-Holstein	-	-	-	-	-0.057	0.021	-0.021	0.018
Hamburg	-	-	-	-	0.023	0.022	0.060	0.022
Niedersachsen	-	-	-	-	-0.083	0.018	0.014	0.015
Bremen	-	-	-	-	-0.081	0.030	-0.018	0.028
Nordrhein-Westfalen	-	-	-	-	-0.046	0.016	0.022	0.014
Hessen	-	-	-	-	0.000	0.018	0.021	0.016
Rheinland-Pfalz	-	-	-	-	-0.044	0.018	-0.002	0.015
Baden-Würtemberg	-	-	-	-	0.000	0.016	0.057	0.015
Bayern	-	-	-	-	-0.006	0.016	0.026	0.014
1992	0.303	0.015	0.241	0.014	0.060	0.014	0.064	0.010
1993	0.518	0.016	0.437	0.015	0.110	0.015	0.111	0.010
1994	0.628	0.016	0.562	0.014	0.125	0.014	0.125	0.010
1994	0.704	0.017	0.635	0.015	0.169	0.015	0.161	0.011
1994	0.736	0.016	0.675	0.015	0.190	0.013	0.187	0.010
1994	0.762	0.016	0.711	0.016	0.190	0.013	0.201	0.010
1994	0.775	0.017	0.726	0.016	0.208	0.014	0.216	0.010
1994	0.777	0.016	0.713	0.016	0.190	0.013	0.210	0.010
Constant	-0.050	0.178	1.428	0.070	1.156	0.062	1.552	0.075
λ	0.190	0.047	-0.019	0.012	-0.013	0.015	-0.123	0.024
Log-Likelihood	-1,414.42				-1,414.24			

Dependent variable: ln(wage).

Chapter 3

Unemployment Dynamics among Migrants and Natives[1]

3.1 Introduction

As part of a strategy to foster growth, migrants have been identified as a target group within the European Union strategy to raise employment levels (Zimmermann, 2005). Of concern is that unemployment is typically very high among migrants with a tendency to rise over time. For instance, since the early 1970s, the unemployment rates of natives and migrants in Germany bifurcate. In 2005, the average share of unemployed migrants has been 25.2% in comparison to the much lower 11.9% among natives (Bundesagentur für Arbeit, 2006). This higher rate of unemployment could derive from a higher risk of becoming unemployed, i.e. a higher frequency of unemployment spells or shorter periods of employment, as well as from a lower probability of leaving unemployment, i.e. a longer duration of unemployment spells. It is important to understand why individuals leave and reenter unemployment, and whether these processes differ between natives and migrants.

Germany can be considered to be a interesting case to investigate the duration of unemployment and employment issue in the context of native-migrant differences. For long, Germany receives the largest migratory flows in the European Union. Nowadays, nearly 20% of the people living in Germany (or 15

[1]The following analysis is based on joint work with Klaus F. Zimmermann (Uhlendorff and Zimmermann, 2006).

million people) are from families with a migration background, one third of the children in the Kindergarden age are from migration families. Hence, the assimilation of immigrants into the German economic system has been subject to much research. For a recent overview of those attempts see Bauer, Dietz, Zimmermann, and Zwintz (2005). The previous literature largely deals with differences in wages between natives and migrants and their assimilation over time. Examples are Dustmann (1993), Schmidt (1997), Fertig and Schurer (2006) and Lang (2005). While the initial earnings gap between immigrants and native workers narrows over time in the U.S. and the U.K. (Borjas, 1994), the evidence is mixed for Germany. There exist, however, only very few studies dealing with unemployment experiences of migrants in Germany. One is the early contribution by Mühleisen and Zimmermann (1994), who deal with the frequency of unemployment among natives and migrants and apply simulated probit estimators. Their results indicate that guest workers do not have a higher risk of becoming unemployed after controlling for observed and unobserved heterogeneity.[2] To our knowledge only Kogan (2004) investigates unemployment and employment durations of migrants.

Hence, this study investigates both sources of higher unemployment rates, unemployment duration and employment stability. The two processes are determined by observed and unobserved characteristics and it is reasonable that the unobserved characteristics influencing both durations are not independent from each other. Therefore we are interested to estimate unemployment and subsequent employment duration models simultaneously and allow for correlation between unobserved terms. Departing from Kogan (2004) we concentrate on immigrants from five guestworker countries (Greece, Italy, Spain, Turkey, Ex-Yugoslavia), take the potential dependence of the two durations into account and analyze subsequent employment duration, conditional on previous unemployment.

Section 3.2 explains the panel data used. Section 3.3 outlines our novel

[2]Additional earlier studies are Cramer (1984) and Bender and Karr (1993), both analyzing the probability of being unemployed using the "Beschäftigtenstatistik", and Winkelmann and Zimmermann (1993) analyzing the frequency of job changes and unemployment spells between 1974 and 1984 applying count data models and using the retrospective data of the SOEP.

econometric approach. Section 3.4 presents the empirical results and section 3.5 summarizes and discusses the implications for economic policy.

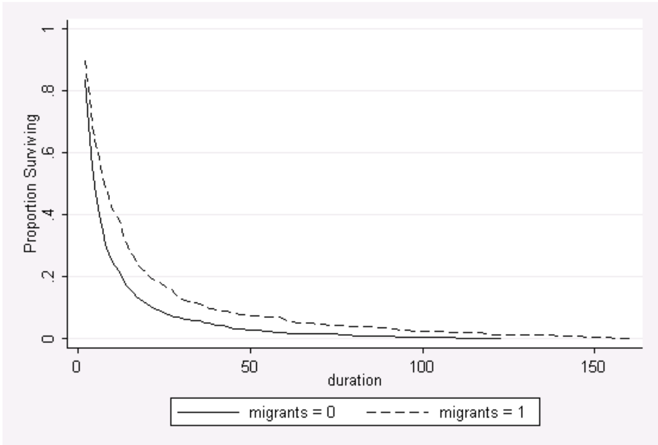
3.2 Data

This study uses data from the German Socio-Economic Panel (SOEP). The annual survey started 1984 in West Germany with a sample of about 5,900 households, 1,400 of them with a household head from one of the main guestworker ethnicities: Turks, Greeks, Italians, ex-Yugoslavians or Spaniards. These migrant groups were over-sampled. 1994/1995 a new migrant sample started consisting of households in which at least one household member migrated to Germany within the last ten years. For a detailed description of the survey see Haisken-DeNew and Frick (2005).

Our sample focuses on migrants from the guestworker countries. This enables us to analyze the labor market assimilation of a group of relatively similar individuals, but also enables us to study potential differences between sufficiently large groups of ethnicities in our sample. We concentrate also on west German natives and migrants only. In east Germany, the share of immigrants is very low and we observe in the SOEP only a few unemployed migrants. Furthermore, east and west German labor markets still exhibit large differences which would attract attention away from our major research topic. However, we include individuals with a migratory background born in Germany if they have have not taken the German citizenship.

Every wave contains retrospective monthly information about the individual employment status of the previous calendar year. We distinguish three categories: employment, unemployment and out of the labor force. The category employment includes full time and half time employment. Out of the labor force includes being in retirement, parental leave, school, university, vocational training and military service. We exclude individuals younger than 20 and older than 55 years, the latter because of special early retirement regulations in Germany during our observation period. Spells of individuals who become 56 years old during the

Figure 3.1: Survivor Functions in Unemployment, Men



Source: SOEP, waves 1984-2004.

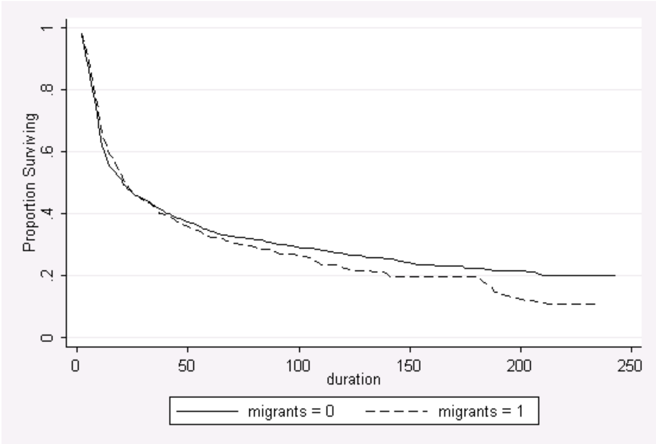
observation period are right-censored at the beginning of the year of the fifty-sixth birthday. Since we only have the information of the year of birth, we right-censor spells at the beginning of the corresponding year and not at the month of the birthday.

Only individuals entering unemployment between 1983 and 2003 are included in the analysis. Our sample consists of unemployment spells and subsequent employment spells. Note that individuals with employment spells enter our sample only if we observe a transition from unemployment to employment. Individuals who are unemployed several times between 1983 and 2003 are in our sample with several spells of unemployment and of subsequent employment. A transition from unemployment to employment is defined as a situation where the employment spell begins at the latest two months after the unemployment spell ends; a transition from employment to unemployment is defined similarly.

The data set used in this paper consists of 4,368 unemployment and 3,080 employment spells of 2,427 individuals. Among the natives, we have 3,111 unemployment spells and 2,204 employment spells. Among the migrants, there are 1,257 unemployment spells and 876 employment spells.

In Figures 3.1 and 3.2 product-limit estimates of the survival functions for both groups, migrants and natives, are presented. They refer to the survivor

Figure 3.2: Survivor Functions in Employment, Men



Source: SOEP, waves 1984-2004.

probability in unemployment and employment, independent of the destination state. At each point in time the share of individuals who are still unemployed is higher for immigrants than for natives. The log-rank test for equality of survivor functions as well as the likelihood-ratio test statistic of homogeneity indicate that the survival functions of both groups differ significantly from each other. For the duration of employment spells the log-rank test indicates no significant difference while the likelihood-ratio test indicates a difference between the two groups at a 10% level. Natives and migrants seem to differ mainly in their unemployment duration and seem to be more similar in their employment duration. Both test-statistics follow a χ^2-distribution with one degree of freedom. The values of the test statistics are 232.55 for the likelihood-ratio test and 125.11 for the log-rank test with respect to the unemployment duration. For the employment duration the corresponding values of the test statistics are 3.12 and 0.03, respectively.

The length of unemployment in our sample ranges from 1 to 160 months, the length of employment spells ranges from 1 to 242 months. Corresponding to the difference in survivor functions in unemployment the average observed length of unemployment spells differs between migrants and natives, see Table 3.1. The observed mean length to a transition to employment is 5.8 months for natives and 8.8 months for migrants, the corresponding mean length for transitions out

Table 3.1: Length and Destination states,unemployment spells

Destination state	Freq.	Percent	Average Length
Natives			
Right censored	371	11.5	12.8
Transitions out of Labor Force	531	17.1	9.3
Transitions to employment	2,209	71.0	5.8
Migrants			
Right censored	208	16.6	19.1
Transitions out of Labor Force	173	13.8	17.4
Transitions to employment	876	69.7	8.8
Total	4,368		8.5

Table 3.2: Length and Destination states,employment spells

Destination state	Freq.	Percent	Average Length
Natives			
Right censored	885	40.2	55.9
Transitions out of Labor Force	301	13.7	31.0
Transitions to unemployment	1,018	46.2	15.0
Migrants			
Right censored	340	38.8	47.9
Transitions out of Labor Force	97	11.1	35.4
Transitions to unemployment	439	50.1	17.6
Total	3,080		32.9

of labor force is 9.3 months for natives and 17.4 months for migrants. Note that these observed mean lengths do not take the censored spells and competing risks into account, which are provided separately in Table 3.1, but nevertheless provide us with a good description of the data set. Around 70% of the observed unemployment spells end due to a transition into employment. Table 3.2 reports the average lengths of employment spells. The observed average length of employment spells exceeds the length of unemployment spells and the differences between migrants and natives are not as striking as in the case of unemployment, which corresponds to the similar survivor functions.

Descriptive statistics of covariables are documented in Table 3.3 separated for natives and migrants and unemployment and employment spells, respectively. Many of those are fixed, but covariables age, marriage status, children in the household, GNP and local unemployment rate are time-variant and they are updated on a yearly level. To control for seasonal effects within the year, dummies for the quarter in which the spell begins are included (first quarter to fourth quarter). For both natives and migrants most unemployment spells begin in the first quarter of the respective year, i.e. between January and March. On average, native and migrant men have the same age (around 33 years). We include educational dummies for the dual-system apprenticeship, additional vocational training and a university degree. Natives have on average a higher education, while more migrants are married and live together with their spouse and they have more often children. Another variable used is disability or handicap. To be disabled means that the individual responds positively to the question whether he is officially registered to have a reduced capacity for work or of being severely disabled. When unemployed, natives and migrants have a 6% share of disabled persons. This share decreases to 5% among natives and to 3% among migrants in subsequent employment spells.

The previous unemployment duration is higher among migrants if they enter a new employment spell. The mean of the local unemployment rate is slightly higher for natives than for migrants when they enter unemployment or employment. In addition to the regional unemployment rate we include the yearly growth rate of the Gross National Product (GNP) in west Germany, which is slightly higher for migrants than for natives, indicating that the migrants in our sample enter unemployment more often in years with relatively high growth rates.

With respect to unemployment spells 42% of those observations are from migrants born in Turkey, 19% are from Ex-Yugoslavians, 15% from Italians, 6% from Greeks, and 3% from Spainiards. 16% of the observations are from migrants who are born in Germany and are, therefore, members of the so called second generation, and around 40% of this second generation have the Turkish citizenship. The German active recruitment policy for guest workers was terminated in the end of 1973 and the following period was characterized by migration through family reunification. Around 40% of the immigrants in our sample arrived before 1974. With respect to the first month of each spell the observed characteristics are similar distributed among the unemployment and employment spells.

Natives and migrants differ with respect to several observable characteristics. These differences could explain differences in the duration of unemployment

Table 3.3: Descriptive Statistics

	Natives		Migrants	
	Unemployment	Employment	Unemployment	Employment
Quarter 1	0.31	0.28	0.33	0.31
Quarter 2	0.18	0.33	0.21	0.31
Quarter 3	0.24	0.22	0.21	0.23
Quarter 4	0.27	0.18	0.25	0.15
Age	32.82 (9.98)	33.25 (9.40)	33.43 (10.71)	32.76 (10.01)
Apprenticeship	0.51	0.52	0.17	0.19
Vocational training	0.14	0.15	0.21	0.19
University	0.12	0.13	0.04	0.03
Married	0.41	0.44	0.62	0.61
Children aged < 4	0.15	0.16	0.26	0.27
Children aged $\geq 4 < 15$	0.24	0.25	0.42	0.43
Handicap	0.06	0.05	0.06	0.03
Previous unemp. duration	-	5.75 (7.32)	-	8.81 (11.92)
Local unemployment rate	9.16 (2.49)	9.13 (2.49)	8.61 (2.68)	8.52 (2.63)
GNP	1.67 (1.65)	1.85 (1.59)	1.91 (1.70)	2.08 (1.61)
Greece	-	-	0.06	0.06
Italy	-	-	0.15	0.16
Spain	-	-	0.03	0.03
Turkey	-	-	0.42	0.39
Ex-Yugoslavia	-	-	0.19	0.19
Second Generation	-	-	0.16	0.17
Second Generation Turkey	-	-	0.07	0.07
Migration before 1974	-	-	0.42	0.39
Number of observations	3,111	2,204	1,257	876

Source: SOEP, numbers refer to first month of each spell, standard deviations in parentheses.

and employment. In addition to that the two groups could differ with respect to unobservable characteristics. This needs to be distinguished from the status effect of being a migrant which could also cause a longer duration of unemployment, e.g. due to discrimination or difficulties with the native language. To analyze these differences in detail we apply econometric methods introduced in the following section.

3.3 Econometric Approach

In this study, we are interested in the duration of and the interdependence between the states unemployment and employment. The process of leaving unemployment for paid labor and the duration of the subsequent employment spell can appropriately be modelled by a multivariate hazard rate model. However, on the labor market we can distinguish three states: unemployment, employment and out of the labor force. There exist two potential levels of dependence via correlated error-terms: Correlations between competing risks and correlations between the duration in different states. The category "out of the labor force" unifies several different categories like early retirement, military service and education. Due to the heterogeneity within this category and the small number of males being in the main working age and not working or searching for work we take the category "out of the labor force" as an independent competing risk into account, i.e. we treat transitions out of the labor force as right-censored, and we do not estimate its duration. Therefore our model ends up in a bivariate hazard rate model consisting of two potentially correlated states, unemployment and employment. For a discussion of multivariate mixed proportional hazard models see van den Berg (2001). According to the type of data being used here - monthly interval-censored observation of the status - discrete time hazard rate models have to be applied (see for example Han and Hausman, 1990, Narendranathan and Stewart, 1993, or Jenkins, 2004).

In the context of employment dynamics, the initial conditions problem arises, because the initial (inflow-) sample of unemployed individuals cannot be assumed to be random, see e.g. Heckman (1981b). This initial conditions problem can be ignored in this study, because we are interested in the subpopulation consisting of individuals entering unemployment. Therefore, the results have to be interpreted with respect to this subpopulation.

The duration of unemployment and employment is generated by a continuous time process. The overall hazard rate $\lambda_s(t)$ for each state s is defined as the limit of the conditional probability for the ending of a spell in interval $[t, t + \triangle t[$ given that no transition occurred before the start of this interval:

$$\lambda_s(t) = \lim_{\triangle t \to 0} \frac{P(t \leq T_s \leq t + \triangle t \mid T_s \geq t)}{\triangle t} \tag{3.1}$$

where T_s denotes the length of a spell. T_s is assumed to be a continuous, non-negative random variable. We assume proportional transition rates with covariates causing proportional shifts of a so-called baseline transition rate and interval constant covariates. For unemployment spells ($s = u$), as well as for employment spells ($s = e$), there exist several potential destination states. Two potential destination states d are considered reflecting transitions into employment and into unemployment ($d = 1$), respectively, and transitions out of labor force ($d = 2$).

$$\lambda_s(t \mid x_i(t), \eta_i) = \sum_{d=1}^{2} \lambda_{sd}(t \mid x_i(t), \eta_{isd}); \quad s = \{u, e\} \tag{3.2}$$

with the hazard rate from state s to destination state d corresponding to

$$\lambda_{sd}(t \mid x_i(t), \eta_{isd}) = \lambda_{0sd}(t) \exp(x_i(t)\beta_{sd} + \eta_{isd}). \tag{3.3}$$

$\lambda_{0sd}(t)$ denotes the state and destination specific baseline transition rate, $x_i(t)$ an individual time variant row vector of covariates for individual i, β_{sd} a column vector of parameters, η_{isd} a time invariant individual unobserved term that varies with state and destination. The unobserved heterogeneity η_i is assumed to be independent of the observed individual characteristics.

We observe the duration of unemployment and employment in monthly intervals. This implies that instead of continuous levels of $x_i(t)$ their interval specific levels have to be taken into account. Assumed that the time axis is divided into intervals of unit length, a given spell consists of a number of j intervals, in the following referred to as subspells. The interval specific levels of $x_i(t)$ and the observed interval baseline hazard $\lambda_{0sd}(t)$ for the $k-th$ subspell are denoted as x_{ik} and $h_{0sd}(k)$.

For interval-censored data with underlying continuous time processes the state-specific survivor function is given by:

$$\begin{aligned} S_s(j|x_i, \eta_i) = \quad & \exp\Big(-\sum_{d=1}^{2}\sum_{k=1}^{j}\exp(x_{ik}\beta_{sd} + h_{0sd}(k) + \eta_{isd})\Big); \\ = \quad & S_{s1}(j)S_{s2}(j); \quad S_{sd}(j) = \exp\Big(-\sum_{k=1}^{j}\exp(x_{ik}\beta_{sd} + h_{0sd}(k) + \eta_{isd})\Big); \\ & h_{0sd}(k) = \ln\Big(\int_{t_{k-1}}^{t_k}\lambda_{0sd}(\tau)d\tau\Big). \end{aligned} \tag{3.4}$$

The survivor function $S_s(j)$ describes the probability that a spell lasts at least j intervals. The h_0 parameters are capturing the duration dependence of the baseline transition function and correspond to the log of the integrated destination-specific baseline hazard rate. The survival function is separable into two destination-specific parts.

In principle, the transitions could occur at any time during the observed intervals. In our approach we assume that transitions can only occur at the boundaries of the intervals (for a similar approach see e.g. Narendranathan and Stewart, 1993). This is a reasonable approximation because new employment is often taken up at the beginning of a month. In the absence of a correlation between the destination specific unobserved heterogeneity terms this leads to two independent risk-specific hazard rates, both following a complementary log-log form and ends up in a separable likelihood with respect to the two independent risks. This implies that transitions from unemployment into employment are independent from transitions out of the labor force and that transitions from employment into unemployment are also independent from transitions out of the labor force, given the observed characteristics. Therefore one can estimate the transition processes, described by transition probabilities $h_{sd}(j)$, separately within each state.

The probabilities $h_{ue}(j)$ and $h_{eu}(j)$ of a transition from unemployment to employment and from employment to unemployment in interval j, respectively, correspond to:

$$
\begin{aligned}
h_{ue}(j|x_i, \eta_i) &= 1 - \exp(-\exp(x_{ij}\beta_{ue} + h_{0ue}(j) + \eta_{iue}) \\
h_{eu}(j|x_i, \eta_i) &= 1 - \exp(-\exp(x_{ij}\beta_{eu} + h_{0eu}(j) + \eta_{ieu})
\end{aligned} \tag{3.5}
$$

This study focusses on the transitions from unemployment to employment and the probability of reentering unemployment. The state specific unobserved heterogeneity components of these transition processes are allowed to be correlated across the two states. Therefore both processes, transitions from unemployment to employment and the process of reentering unemployment again have to be estimated jointly. Transitions out of the labor force enter the estimation as right-censored spells. The joint estimation is important because there is no reason to believe that unobserved characteristics determining the duration of unemployment are independent from unobserved characteristics influencing subsequent employment stability. Ignoring this could create a sample selection problem and thereby yield biased estimates. For a similar argument in the context of experimental data on training and the selection into subsequent employment spells see Ham and LaLonde (1996).

η_{ue} is the unobserved heterogeneity influencing the transition process from unemployment to employment, while the unobserved term η_{eu} effects employment stability. Following Heckman and Singer (1984) these unobserved terms or random intercepts are assumed to follow a discrete probability distribution with a finite number of mass points η_{sd}^m, $m = (1, ..., M)$.

The indicators δ_u and δ_e take on the value 1 if a transition to employment or to unemployment, respectively, is observed and zero otherwise. The likelihood contribution of an unemployment spell of j_u intervals and a subsequent employment spell of j_e intervals for a given x_i, η_{iue} and η_{ieu} is:

$$\begin{aligned} l(x_i, \eta_{iue}, \eta_{ieu}) = \quad & S_{u1}(j_u - 1|x_i, \eta_{iue}) h_{ue}(j_u|x_i, \eta_{iue})^{\delta_u} (1 - h_{ue}(j_u|x_i, \eta_{iue}))^{(1-\delta_u)} \\ & S_{e1}(j_e - 1|x_i, \eta_{ieu})^{\delta_u} h_{eu}(j_e|x_i, \eta_{ieu})^{\delta_e \delta_u} (1 - h_{eu}(j_e|x_i, \eta_{ieu}))^{(1-\delta_e)\delta_u}. \end{aligned}$$

The unobserved heterogeneity is assumed to follow a multivariate distribution $G(\eta_{ue}, \eta_{eu})$ with a finite number of points of support. Each term has three points of support. This results in 9 points of support for G: $(\eta^1_{ue}, \eta^1_{eu})$, $(\eta^1_{ue}, \eta^2_{eu})$, ... and $(\eta^3_{ue}, \eta^3_{eu})$. For each of these combinations there exists a probability or a share of individuals having these values of unobserved heterogeneity. For a similar modelling of unobserved heterogeneity with two points of support for each random term see e.g. Stevens (1999) in the context of income poverty duration or Belzil (2001) in the context of unemployment and subsequent employment duration. The likelihood contribution of an unemployment and a subsequent employment spell for a given x_i but unknown η_{iue} and η_{ieu} can be written as

$$\begin{aligned} l(x_i, \eta_{ue}, \eta_{eu}) = \quad & \pi_1 * l(x_i, \eta^1_{ue}, \eta^1_{eu}) + \pi_2 * l(x_i, \eta^1_{ue}, \eta^2_{eu}) + \pi_3 * l(x_i, \eta^1_{ue}, \eta^3_{eu}) + \\ & \pi_4 * l(x_i, \eta^2_{ue}, \eta^1_{eu}) + \pi_5 * l(x_i, \eta^2_{ue}, \eta^2_{eu}) + \pi_6 * l(x_i, \eta^2_{ue}, \eta^3_{eu}) + \\ & \pi_7 * l(x_i, \eta^3_{ue}, \eta^1_{eu}) + \pi_8 * l(x_i, \eta^3_{ue}, \eta^2_{eu}) + \pi_9 * l(x_i, \eta^3_{ue}, \eta^3_{eu}) \end{aligned} \quad (3.6)$$

For the estimation procedure the probabilities π_l are specified as logistic probabilities to ensure that the probabilities vary between 0 and 1 and add up to 1.[3]

$$\pi_l = \frac{\exp(p_l)}{\sum_{r=1}^{9} \exp(p_r)}, \quad l = 1, ..., 9, \quad \sum_{r=1}^{9} \pi_r = 1 \quad (3.7)$$

As the hazard rates contain a constant term, for identification reasons one of the mass points of each unobserved heterogeneity term η_{ue} and η_{eu} and one of the parameters p_r are normalized to 0.

In the data we observe several spells for some individuals. We assume that

[3] The model has been programmed in Stata version 8.2

the unobserved heterogeneity terms are constant for each individual i. Therefore the unobserved heterogeneity has to be integrated out over all Q_i spells of one individual. For a similar treatment of repeated spells per individual see e.g. Steiner (2001) or Roed and Zhang (2005).

The sample likelihood is given by

$$L = \prod_{i=1}^{n} \sum_{r=1}^{9} \pi_r \prod_{q=1}^{Q_i} l_q(x_i, \eta_{ue}, \eta_{eu}) \tag{3.8}$$

3.4 Results

We estimate a bivariate discrete time hazard rate model with jointly distributed unobserved heterogeneity. The coefficients can be interpreted with respect to the underlying continuous time proportional hazard rates. Compared to the model without unobserved heterogeneity, the inclusion of unobserved heterogeneity does significantly improve the model fit. The results of the models without and with unobserved heterogeneity are reported in Tables 3.4 and 3.5, respectively. All estimated mass-points are significantly different from 0. The coefficients indicate that there exist three groups in both processes which differ significantly from each other with respect to the baseline hazard rate. The point estimates suggest that the hazard rate from unemployment to employment is reduced by 69% for one group and increased by a factor of 2.4 for another group. With respect to the probability of staying employed one group has a 75% reduced risk of leaving employment and this probability is nearly 5 times higher for another group.

Two of the nine probabilities describing the distribution G of the unobserved heterogeneity converge to zero. This indicates that two combinations of unobserved heterogeneity terms do not exist. In the estimation procedure we set these points of support to zero and a distribution with seven points of support remains. The distribution is shown in Table 3.6.

Table 3.4: Estimation results without unobserved heterogeneity

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1				Model 2				Model 3			
	Unemployment to Employment		Employment to Unemployment		Unemployment to Employment		Employment to Unemployment		Unemployment to Employment		Employment to Unemployment	
Months 4-6	-0.16**	0.05	0.10	0.09	-0.16**	0.05	0.10	0.09	-0.16**	0.05	0.10	0.09
Months 7-12	-0.55**	0.05	0.56**	0.08	-0.55**	0.05	0.56**	0.08	-0.54**	0.05	0.56**	0.08
Months 13-18	-0.78**	0.08	-0.38**	0.11	-0.78**	0.08	-0.38**	0.11	-0.76**	0.08	-0.38**	0.11
Months 19+	-1.57**	0.09	-0.93**	0.11	-1.57**	0.09	-0.93**	0.11	-1.54**	0.09	-0.94**	0.11
Quarter 1	0.01	0.05	-0.03	0.09	0.01	0.05	-0.04	0.09	0.01	0.05	-0.03	0.09
Quarter 2	-0.24**	0.06	0.21*	0.08	-0.23**	0.06	0.21*	0.08	-0.22**	0.06	0.21*	0.08
Quarter 3	-0.17**	0.05	0.05	0.09	-0.17**	0.05	0.05	0.09	-0.18**	0.05	0.05	0.09
December	0.21	0.06	0.87**	0.07	0.21**	0.06	0.87**	0.07	0.21**	0.06	0.88**	0.07
Age	0.04**	0.01	-0.14**	0.02	0.05**	0.01	-0.14**	0.02	0.04**	0.01	-0.14**	0.02
Age squared $*10^{-2}$	-0.10**	0.02	0.21**	0.02	-0.10**	0.02	0.21**	0.02	-0.10**	0.02	0.21**	0.02
Married	0.30**	0.05	-0.31**	0.07	0.30**	0.05	-0.31**	0.07	0.32**	0.05	-0.32**	0.07
Children aged < 4	-0.09	0.05	0.15*	0.07	-0.09	0.05	0.15*	0.07	-0.08	0.05	0.14	0.07
Children aged $\geq 4 < 15$	-0.05	0.04	0.22**	0.06	-0.05	0.04	0.22**	0.06	-0.04	0.04	0.21**	0.06
Apprenticeship	0.24**	0.05	-0.29**	0.07	0.23**	0.05	-0.28**	0.07	0.22**	0.05	-0.29**	0.07
Vocational training	0.23**	0.06	-0.40**	0.0	0.23**	0.07	-0.40**	0.08	0.22**	0.06	-0.41**	0.086
University	0.49**	0.07	-0.65**	0.11	0.49**	0.07	-0.65**	0.11	0.49**	0.07	-0.66**	0.11
Handicap	-0.60**	0.09	0.05	0.13	-0.60**	0.09	0.05	0.13	-0.61**	0.09	0.05	0.13
Local unemp. rate	-0.05**	0.01	0.02	0.01	-0.06**	0.01	0.02	0.01	-0.05**	0.01	0.02	0.01
GNP	0.08**	0.01	-0.04**	0.02	0.08**	0.01	-0.04**	0.02	0.08**	0.01	-0.04**	0.02
Previous unemp. duration $*10^{-2}$	-	-	-0.47	0.53	-	-	-0.47	0.53	-	-	-0.59	0.54
Prev. un. dur. squared $*10^{-3}$	-	-	0.02	0.07	-	-	0.02	0.07	-	-	0.03	0.07
Migrant	-0.34**	0.05	-0.09	0.07	-0.38**	0.05	-0.07	0.07	-	-	-	-
Second generation	-	-	-	-	-0.22*	0.09	-0.14	0.13	-0.06	0.11	-0.11	0.16
Second generation Turkey	-	-	-	-	-	-	-	-	-0.43**	0.14	-0.15	0.21
Turkey	-	-	-	-	-	-	-	-	-0.54**	0.07	0.07	0.10
Spain	-	-	-	-	-	-	-	-	0.23	0.20	-0.29	0.31
Italy	-	-	-	-	-	-	-	-	-0.16	0.10	0.05	0.14
Ex-Yugoslavia	-	-	-	-	-	-	-	-	-0.17	0.09	0.01	0.14
Greece	-	-	-	-	-	-	-	-	-0.31*	0.15	-0.43	0.25
Migration before 1974	-	-	-	-	-	-	-	-	-0.06	0.08	-0.15	0.11
Constant	-1.99**	0.25	-1.38**	0.30	-2.02**	0.25	-1.37**	0.30	-2.00**	0.26	-1.44**	0.30
Log-Likelihood	-16,705.90				-16,703.76				-16,682.13			

Observations: 4,368 unemployment spells (37,174 months), 3,080 employment spells (101,464 months).
*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Table 3.5: Estimation results with unobserved heterogeneity

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1				Model 2				Model 3			
	Unemployment to Employment		Employment to Unemployment		Unemployment to Employment		Employment to Unemployment		Unemployment to Employment		Employment to Unemployment	
Months 4-6	0.13*	0.05	0.18	0.09	0.12*	0.05	0.18	0.09	0.13*	0.05	0.18	0.09
Months 7-12	-0.09	0.07	0.79**	0.08	-0.09	0.07	0.79**	0.08	-0.07	0.07	0.79**	0.08
Months 13-18	-0.16	0.09	-0.02	0.12	-0.16	0.09	-0.02	0.12	-0.14	0.09	-0.02	0.12
Months 19+	-0.73**	0.11	-0.39**	0.12	-0.74**	0.11	-0.39**	0.12	-0.70**	0.12	-0.39**	0.12
Quarter 1	0.07	0.06	-0.07	0.10	0.07	0.06	-0.07	0.10	0.08	0.06	-0.07	0.10
Quarter 2	-0.16*	0.07	0.14	0.10	-0.16*	0.07	0.14	0.10	-0.14*	0.07	0.14	0.10
Quarter 3	-0.07	0.07	0.08	0.11	-0.07	0.07	0.08	0.11	-0.07	0.07	0.09	0.11
December	0.20**	0.06	0.90**	0.07	0.20**	0.06	0.90**	0.07	0.20**	0.06	0.90**	0.07
Age	0.06**	0.02	-0.16**	0.02	0.06**	0.02	-0.16**	0.02	0.05**	0.02	-0.16**	0.02
Age squared $*10^{-2}$	-0.13**	0.03	0.22**	0.02	-0.13**	0.03	0.23**	0.02	-0.13**	0.03	0.23**	0.02
Married	0.32**	0.06	-0.30**	0.08	0.33**	0.06	-0.30**	0.08	0.34**	0.06	-0.30**	0.08
Children aged < 4	-0.14*	0.06	0.10	0.08	-0.14*	0.06	0.10	0.08	-0.13*	0.06	0.10	0.08
Children aged $\geq 4 < 15$	-0.04	0.06	0.21**	0.07	-0.04	0.06	0.21*	0.07	-0.02	0.06	0.21**	0.07
Apprenticeship	0.32**	0.07	-0.29**	0.08	0.32**	0.07	-0.28**	0.08	0.30**	0.06	-0.28**	0.08
Vocational training	0.40**	0.08	-0.47**	0.10	0.40**	0.08	-0.48**	0.10	0.39**	0.08	-0.48**	0.11
University	0.65**	0.10	-0.81**	0.14	0.65**	0.10	-0.80**	0.14	0.64**	0.09	-0.80**	0.14
Handicap	-0.71**	0.11	0.11	0.16	-0.71**	0.11	0.11	0.16	-0.74**	0.11	0.12	0.16
Local unemp. rate	-0.06**	0.01	0.04**	0.01	-0.06**	0.01	0.04**	0.01	-0.06**	0.01	0.04**	0.01
GNP	0.10**	0.01	-0.04*	0.02	0.10**	0.01	-0.04*	0.02	0.10**	0.01	-0.04*	0.02
Previous unemp. duration $*10^{-2}$	-	-	-0.25	0.72	-	-	-0.24	0.72	-	-	-0.15	0.73
Prev. un. dur. squared $*10^{-3}$	-	-	0.01	0.09	-	-	0.01	0.09	-	-	0.00	0.10
Migrant	-0.40**	0.07	-0.08	0.08	-0.44**	0.07	-0.05	0.09	-	-	-	-
Second generation	-	-	-	-	-0.31*	0.12	-0.22	0.18	-0.12	0.15	-0.23	0.21
Second generation Turkey	-	-	-	-	-	-	-	-	-0.58**	0.18	-0.18	0.30
Turkey	-	-	-	-	-	-	-	-	-0.67**	0.10	-0.05	0.12
Spain	-	-	-	-	-	-	-	-	0.26	0.27	-0.40	0.44
Italy	-	-	-	-	-	-	-	-	-0.08	0.14	0.08	0.17
Ex-Yugoslavia	-	-	-	-	-	-	-	-	-0.13	0.13	-0.10	0.17
Greece	-	-	-	-	-	-	-	-	-0.35	0.20	-0.42	0.31
Migration before 1974	-	-	-	-	-	-	-	-	-0.14	0.10	0.04	0.14
η_{ue2}, η_{eu2}	-1.16**	0.14	-1.37**	0.10	-1.15**	0.14	-1.38**	0.10	-1.18**	0.15	-1.37**	0.10
η_{ue3}, η_{eu3}	1.19**	0.11	1.57**	0.18	1.18**	0.11	1.58**	0.18	1.21**	0.10	1.59**	0.18
Constant	-2.33**	0.37	-0.73*	0.35	-2.35**	0.37	-0.72*	0.35	-2.33**	0.37	-0.75*	0.35
Log-Likelihood	-16,485.88				-16,484.45				-16,465.87			

Unobserved heterogeneity is assumed to follow a non parametric distribution. For both processes 2 mass points are freely estimated.

Observations: 4,368 unemployment spells (37,174 months), 3,080 employment spells (101,464 months).

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Table 3.6: Distribution of unobserved heterogeneity, Model 3

	Prob.	Std. Err.	Prob.	Std. Err.
P(η^1_{ue}, η^1_{eu}), P(η^1_{ue}, η^2_{eu})	14.1	3.7	45.4	6.3
P(η^1_{ue}, η^3_{eu}), P(η^2_{ue}, η^1_{eu})	1.3	0.4	9.1	3.8
P(η^2_{ue}, η^2_{eu}), P(η^2_{ue}, η^3_{eu})	12.9	6.6	0	-
P(η^3_{ue}, η^1_{eu}), P(η^3_{ue}, η^2_{eu})	7.1	1.6	10.1	3.3
P(η^3_{ue}, η^3_{eu})	0	-		

The standard errors of the probabilities are derived using the delta method. The results refer to model 3, the distributions of model 1 and model 2 are quite similar.

The largest group (45%) belongs to the base (middle) category with respect to the unemployment duration and remains employed relatively long while the smallest fraction reenters unemployment with a high probability and belongs to the base group with respect to unemployment duration (1.3%). In addition to this model we estimated the processes separately, both with three points of support. Compared to the joint estimation of the duration processes, the Akaike Information Criterion (AIC) indicates that the processes are not independent from each other.[4] However, the increase in the log-likelihood is relatively small (5.8) and the results do not change qualitatively. Alternatively, we estimated our models with two points of support for each random term and found the difference in the log-likelihood between the joint model and the separated models not significant. However, the estimation with three points of support lead to a significant improvement.

We estimated three different models with respect to the included migration variables: in the first model one variable indicating whether a person is a migrant or not is included, in the second model we additionally control for migrants who are born in Germany, and in the third model detailed information about the ethnicity and the year of migration is included. The inclusion of detailed information significantly increases the log-likelihood, indicating that it is important to distinguish between different ethnic groups when analyzing the unemployment dynamics of migrants.

The findings discussed in the sequel are based on the full model capturing

[4] $AIC = -2lnL + 2z$, lnL is the log Likelihood and z the number of parameters, see e.g. Cameron and Trivedi (2005).

unobserved heterogeneity (see model 3 in Table 3.5), since this is the one with the best overall fit. We are particularly interested in the unemployment and employment duration differences between natives and and the migrant groups. We find that Turks and Greeks have a significantly lower hazard rate from unemployment to employment than natives, but the effect for the Greeks is significant only at the 10% level. The point estimate for the Turks suggest that they have a reduction in the hazard rate of around 50% compared to the natives, which is quite substantial. The hazard rates of migrants coming from Italy, Ex-Yugoslavia and Spain do not differ significantly from the hazard rate of native men. Members of the second generation, i.e. children of migrants coming from the guestworker countries, have a 44% lower hazard rate from unemployment to employment if they have a Turkish citizenship; however, there is no difference for second generation migrants with other citizenships. This indicates that job finding difficulties do not disappear for Turkish individuals who were grown up in Germany. Moreover, these results indicate that the economic disadvantages of migrants typically identified in studies on unemployment duration in Germany (see e.g. Steiner, 2001, or Uhlendorff, 2004), are driven by the performance of one ethnic group, the Turks.

Once migrants find a new job, we observe no significant disadvantages of ethnic groups in the employment stability compared to natives. These results suggest that, compared to natives with the same observable and unobservable characteristics, unemployed immigrants do not find less stable jobs but that they need more time to find these jobs. Immigrants who came to Germany before 1974, i.e. before the recruitment policy for guestworkers was terminated, and persons who immigrated afterwards do not differ from each other with respect to both processes.

Our analysis controls for a number of covariables. The results indicate that the probability of finding a job increases in the months 4-6 being unemployed in comparison with the first three months and decreases afterwards. For employed individuals we observe a higher exit rate from jobs to unemployment in the months 7-12, compared to the first half year of employment, and a decreasing exit rate afterwards. Young and old unemployed persons stay longer in unemployment, while young and old employed have less stable jobs than the middle aged. The presence of young children in the household exhibits a higher probability of staying unemployed, while the coefficient of having older children is not significantly different from zero. Small kids do not have a significant impact estimate on the duration of employment, while the presence older children shows a negative impact on job stability. Married men have a higher probability of leaving unemployment as well as a more stable employment.

Higher education protects individuals from unemployment, since all the categories included (apprenticeship, further vocational training and university with no vocational training at all as the reference category) have parameter estimates that strongly indicate a higher probability of leaving unemployment and a more stable employment spell. Individuals with a handicap have a higher risk of staying unemployed, but once they find a job, these jobs are as stable as the jobs of employees without a handicap. For both, the business cycle and the local unemployment rate, we find an impact on unemployment duration but no impact on employment stability. Growth increases the probability to find a job while higher local unemployment rates decrease such a chance.

3.5 Summary and Policy Conclusions

There is much concern in many European countries such as Germany about the very high unemployment rates among migrants and the likely causes. Therefore, this paper has investigated the differences in unemployment dynamics between natives and migrants in Germany to provide evidence about the most relevant factors. Using spell information of the 1984-2004 waves from the German Socio-Economic Panel (SOEP) for men aged between 20 and 55 the analysis is based on an inflow sample into unemployment and the estimation of a bivariate hazard rate model with two states, unemployment and employment.

Two processes are analyzed: Transitions from unemployment to employment and transitions from employment to unemployment. The durations of both states are estimated jointly and the state specific unobserved heterogeneity components are allowed to be correlated across the two states. This is important because there is no reason to believe that unobserved characteristics determining the duration of unemployment are independent from unobserved characteristics influencing subsequent employment stability. Ignoring potential dependence could create a sample selection problem and thereby yield biased estimates. We find some evidence that both processes are not independent from each other, but the results do not change qualitatively compared to a model with uncorrelated unobserved heterogeneity.

The results show that migrants stay longer unemployed than natives, but the probability of leaving unemployment differs strongly with ethnicity. While immigrants from Italy, Ex-Yugoslavia and Spain do not differ from natives, Turkish immigrants have a significantly lower probability of leaving unemployment for a paid job. Moreover Turkish members of the second generation of guestworkers still have a significantly lower probability of leaving unemployment than natives. However, once migrants find a new job, we observe no significant differences in the employment stability compared to natives, independent of the ethnicity.

These results suggest that, compared to natives with the same observable and unobservable characteristics, unemployed immigrants do not find less stable jobs but that they need more time to find these jobs. Predominantly Turks from the first and second generation face the problem of slow integration from unemployment to employment. Therefore, adequate policy measures should concentrate on the job finding process of Turkish migrants to decrease their disadvantages on the labor market.

Chapter 4

From No Pay to Low Pay and Back Again? A Multi-State Model of Low Pay Dynamics

4.1 Introduction

Unskilled and low-skilled individuals have a relatively high risk of unemployment. In this context low pay employment is discussed controversially. In Germany, characterized by an almost continuously rising unemployment rate in the last decades, it is often argued that rising employment rates in the low pay sector could be one solution to overcome the high unemployment rate among low skilled workers. On the other hand low paid jobs are often associated with unstable working careers and a high risk of unemployment. According to this the ongoing public debate in Germany ranges from discussions of the introduction of a minimum wage to the implementation of workfare programs. In this context it is important to know whether low paid jobs are transitory experiences of the working career and stepping stones to better jobs or whether there exists a "low pay - no pay cycle".

In this paper, I analyze low pay and non-employment dynamics of men in west Germany. The focus lies on the extent of true or genuine state dependence in low pay and non-employment. True state dependence describes the fact that being low paid or not employed in one period itself increases the probability of being low paid or not employed in the next period. The knowledge of the state dependence allows to evaluate in how far the employment prospects of low paid individuals differ from not employed and high paid individuals.

The existence of state dependence in employment dynamics can be explained by several factors. Past unemployment may alter preferences, prices or constraints and therefore increase the probability of future unemployment, see e.g. Heckman and Borjas (1980) and Prowse (2005). For example, non-employment may prevent human capital accumulation and lead to a loss of work experience or firms may use unemployment spells as a proxy for unobserved components of ability in their hiring decisions. These effects may be the same for low wage jobs. Being low paid could lead to non-accumulation and deterioration of human capital. Moreover, McCormick (1990) argues that low paid jobs are low-quality jobs and the type of job may be used by firms as an indicator about worker quality. Hence, being low paid could stigmatize employees and may be used as a screening device of employers (Stewart, 2006). The aim of this paper is to examine the extent of true state dependence and to analyze whether low wage jobs have the same or even higher adverse effects on future employment prospects compared to non-employment.

Studies comparing the extent of a low wage sector across countries indicate that there exist wide variations, with the highest incidence of low pay in western Europe measured in the UK and an average incidence in Germany (e.g. European Commission, 2004). Numerous studies exist on low pay dynamics in Europe, e.g. the edited volume of Asplund, Sloane, and Theodossiu (1998) contains several analyses. Descriptive studies about the low pay dynamics indicate that the low pay dynamics have been decreasing in Germany over the last two decades (Rhein, Gartner, and Krug, 2005) and that Germany has the lowest exit probability from low pay to high pay in western Europe (European Commission, 2004). Stewart and Swaffield (1999) have shown that models without potential endogeneity of the initial wage state may lead to biased parameter estimates. This endogeneity of the initial wage state is taken into account only in some of the existing studies (Stewart and Swaffield, 1999, Cappellari, 2002, Sousa-Poza, 2004, among others). Cappellari and Jenkins (2004) additionally allow for potentially endogenous selection into employment and panel attrition. They conclude that 'economic' selection is more important than 'survey' selection. So far, for Germany, there exists no study on low pay dynamics accounting for the endogeneity of the initial state, but several studies on unemployment dynamics. Flaig, Licht, and Steiner (1993) and Mühleisen and Zimmermann (1994) find evidence for state dependence in unemployment and Haan (2005) reports state dependence in employment for married women. These results correspond to the results for other countries. For example Arulampalam, Booth, and Taylor (2000) find state dependence in unemployment for British men, Hyslop (1999) and Michaud and Tatsiramos (2005) find state dependence in employment for married women in the US and in several European countries, respectively, and Prowse (2005) reports state dependence in part- and full-time employment for women in Britain.

As far as I know only two studies investigate the relation between the three labor market states low pay, high pay and unemployment and taking the initial condition problem into account: Cappellari and Jenkins (2004) and Stewart (2006), both using the British Household Panel (BHPS).[1] Cappellari and Jenkins (2004) estimate a multivariate probit model with several endogenous selection processes and find evidence for state dependence and a higher probability of becoming unemployed for low paid and of becoming low paid for unemployed individuals. Stewart (2006) analyzes the transitions into unemployment and takes the previous labor market state into account. In contrast to Cappellari and Jenkins (2004) he makes use of the panel structure of his data set by estimating several dynamic random and fixed effects models including models with autocorrelated error terms, bivariate random effects and GMM estimators. His results do not differ qualitatively between the various methods and are in line with the results of Cappellari and Jenkins.

I extend the approaches of Stewart (2006) and Cappellari and Jenkins (2004) and estimate a dynamic multinomial logit model with random effects. In this

[1]There is evidence of earning losses associated with job loss and unemployment, see e.g. Jacobsen, LaLonde, and Sullivan (1993) and Ketzler (1998) for the U.S. and Gregory and Jukes (2001) for the U.K. However, Wachter and Bender (2006) state that for young German job loosers wage losses fade away within five years.

model, it is possible to differentiate three initial and three destination states instead of two destination states in binary probit models. Therefore, I can model ‘economic’ selection with respect to non-employment as a mutually exclusive state directly in the multinomial model. In addition to that I take the ‘survey’ selection into account by simultaneously modeling the panel attrition similar to Cappellari and Jenkins (2004). In contrast to them I make use of the panel structure of the data and allow for random effects.

Low pay is defined as a relative concept and the models are estimated with two alternative thresholds defined as two-thirds of the median hourly gross wage and the first quintile of the hourly gross wage distribution. All wages above the corresponding threshold are labeled as “high paid”. In my analysis I use data from the German Socio-Economic Panel Study (SOEP) for men aged between 20 and 55.

The results indicate that there exists strong true state dependence in low pay as well as in non-employment. In addition to that there exists a strong link between low pay and no pay. Compared to high paid workers not employed individuals have a higher probability to be low paid in the future and vice versa. Despite this clear evidence for a “low pay - no pay cycle”, compared to non-employment low-wage jobs increase the probability of being employed in the future and low pay seems to lead to higher paid jobs. Thus, there is some evidence that low paid jobs are stepping stones to better jobs and no evidence that being low paid does have any adverse effects on future employment prospects if it is compared with non-employment. However, being low paid goes along with a higher risk of non-employment and a higher probability of being low paid in the future if it is compared to high paid jobs. I find no evidence for the endogeneity of panel attrition. The corresponding correlation coefficients are insignificant and the results do not change compared to the simpler model.

Section 4.2 gives a short description of the data, the low pay definitions and descriptive statistics of the transition probabilities. Section 4.3 outlines the econometric approach, Section 4.4 presents empirical results and section 4.5 concludes.

4.2 Data and Descriptive Statistics

This study uses data from the German Socio-Economic Panel Study (SOEP). The annual survey started in 1984 in west Germany and was extended to include east Germany in 1990. In all panel waves, the head of the household provides information about the household and every household member aged 16 or older provides additional individual information. For a detailed description see Haisken-DeNew and Frick (2005).

Monthly payments may vary due to short-time or overtime working and bonus payments (Sloane and Theodossiu, 1998). Therefore, I measure earnings on an hourly basis accounting for overtime working and excluding bonus-payments, and include full-time, part-time as well as marginal employment. This information is given for the month previous to the interview, hence the labor market information refers to one month in the year.

I define low pay as a relative concept. Individuals whose wage does not exceed a certain relative position in the wage distribution are defined as being low paid. In the literature different low pay cutoffs are used. Stewart and Swaffield (1999) use two thresholds and define low paid employees as persons whose earnings are less than half of the median and whose earnings are less than two-thirds of the median, respectively. Cappellari (2002) uses two thresholds as well and defines the first quintile and the third decile of the wage distribution as low wages. In this study two alternative thresholds are applied: individuals with a gross wage lower (i) than two thirds of the current median hourly earnings and (ii) the first quintile of the wage distribution are defined to be low paid, respectively. In Table 4.1 the hourly low-pay thresholds are presented for the different years in 2000 prices. These low pay thresholds are calculated on an annual basis and refer to all individuals, men and women, not being self-employed, living in west Germany, reporting their working hours and their last monthly wage in the SOEP. The 2/3 median threshold lies in every year below the first quintile threshold and is almost continuously rising, reflecting a real wage growth of the median wage over time. For the first quintile threshold, no clear trend can be observed.

Table 4.1: Low Pay Thresholds 1998-2003 in prices of 2000, Euro

	1998	1999	2000	2001	2002	2003
2/3 median	8.27	8.75	8.88	8.87	8.97	9.23
First quintile	10.47	10.25	10.57	10.58	10.44	11.04

Source: SOEP, weigthed yearly observations

Earning and participation dynamics may differ between men and women and have to be analyzed separately. Therefore I exclude women from the analysis. In addition to that individuals younger than 20 and older than 54 years are excluded from the sample. The first age restriction is motivated by the schooling schemes and the second one by the retirement schemes in Germany. The sample focuses on west Germany. The reason for this is given by the differences in the wage distributions between east and west Germany. Calculating joint thresholds would imply a very small share of low paid individuals in west Germany.[2] Furthermore, east and west German labor markets still exhibit large differences which would draw attention away from the major research topic of this paper. Moreover individuals who are at no interview date during the observation period employed or registered as unemployed are excluded because these individuals have a high probability to be out of the labor force.

I use the SOEP waves 1998 to 2003 for the analysis. An individual enters the sample if the person is within the age restrictions, has finished education, civilian or military service and is not self-employed or in "disabled employment" at any interview date. There exist two possible entry dates: 1998, the first year of observation and the year of the introduction of the "refreshment" sample and 2000, the year of the introduction of the "innovation" sample in the SOEP. Around 45% of all interviewed individuals in 2000 belong to the "innovation" sample.[3] In the regression analysis, described in the next chapter, I control for the entry date 1998 and 2000, respectively, to capture potential differences

[2]For the wage gap between east and west Germany and its development over time see e.g. Görzig, Gornig, and Werwatz (2005).

[3]Both, the refreshment and the innovation samples are supplementary random samples with the aim to stabilize the number of cases in the SOEP (Schupp and Wagner, 2002).

between the two samples with respect to labor market transition processes.[4] An individual leaves the sample in the first year in which it is not possible to observe one of the variables used in the econometric analysis. This could happen by panel attrition or by missing values in the dependent or independent variables. This leads to an unbalanced panel data set with continuously observed years for each individual.

The share of low paid men in west Germany in 1998 is around 6.9% and was increasing to 9.2% in 2003 with respect to the first threshold (2/3 median). The corresponding shares evaluated by the first quintile of the wage distribution are with 16.9% in 1998 and 17.5% in 2003 more stable. Compared to the United Kingdom this is a relatively low rate of low paid employees. For example Stewart and Swaffield (1999) report around 22% of working British men to be low paid in the years 1991-1995 with respect to the first threshold.

Table 4.2 presents the probabilities of being low or high paid in period t, conditional on the pay state in the previous period $t-1$. The unweighted sample consists of pooled year to year transitions between 1998 and 2003 and is restricted to men being employed and reporting wages in at least two waves. The probability of being low paid is much higher for those who have been low paid in the previous year. For the first threshold (2/3 median) around 43% of the low paid individuals stay low paid if they are still employed and less than two percent of the previously high paid individuals are low paid in the next period. The second threshold (first quintile) goes along with a higher state dependence in low paid jobs (47%) and a slightly higher transition probability from high paid to low paid jobs (2.1%).

In Table 4.3 non-employment is additionally taken into account. The pooled sample is restricted to those being not employed or employed with observed wages. Taking the non-employment into account, we still observe a much higher probability of being low paid for those who have been low paid in the previous period compared to previously high paid individuals. The probability of being not em-

[4] Differences between these two cohorts could exist because the 1998 sample mainly consists of individuals who have been taking part in the SOEP for several waves, i.e. the share of individuals with a low probability of attrition is relatively high compared to the 2000 sample.

Table 4.2: Transition Matrix: Low Pay and High Pay

	Threshold 1			Threshold 2		
	Low paid, t	High paid, t	Total	Low paid, t	High paid, t	Total
Low paid, t-1	43.1	57.0	3.5	47.1	53.0	6.0
High paid, t-1	1.6	98.4	96.5	2.1	97.9	94.0
Total	3.0	97.0	100	4.8	95.3	100

Source: SOEP, unweighted pooled sample 1998-2003, n=8,483

Threshold 1: 2/3 of the median, Threshold 2: First quintile

Table 4.3: Transition Matrix: Low Pay, High Pay and Non-Employment

	Threshold 1			
	Non-employment, t	Low paid, t	High paid, t	Total
Non-employment, t-1	72.2	9.3	18.5	7.4
Low paid, t-1	15.2	36.5	48.3	3.7
High paid, t-1	2.4	1.5	96.0	88.9
Total	8.1	3.4	88.5	100
	Threshold 2			
	Non-employment, t	Low paid, t	High paid, t	Total
Non-employment, t-1	72.2	12.1	15.7	7.4
Low paid, t-1	13.3	40.8	45.9	6.2
High paid, t-1	2.2	2.0	95.8	86.4
Total	8.1	5.2	86.8	100

Source: SOEP, unweighted pooled sample 1998-2003, n=9,441

Threshold 1: 2/3 of the median, Threshold 2: First quintile

Table 4.4: Transition Matrix: Low Pay, High Pay, Non-Employment and Attrition

	Threshold 1				
	Non-employment, t	Low paid, t	High paid, t	Attrition, t	Total
Non-employment, t-1	66.0	8.5	16.7	8.6	7.7
Low paid, t-1	14.1	33.9	44.8	7.2	3.8
High paid, t-1	2.3	1.4	90.9	5.4	88.6
Total	7.6	3.2	83.5	5.7	100
	Threshold 1				
	Non-employment, t	Low paid, t	High paid, t	Attrition, t	Total
Non-employment, t-1	66.0	11.1	14.3	8.6	7.7
Low paid, t-1	12.4	38.0	42.8	6.8	6.3
High paid, t-1	2.1	1.9	90.7	5.3	86.1
Total	7.6	4.9	81.8	5.7	100

Source: SOEP, unweighted pooled sample 1998-2003, n=10,010

Threshold 1: 2/3 of the median, Threshold 2: First quintile

Table 4.5: Descriptive Statistics

		Threshold 1		Threshold 2	
	Non-employment	Low paid	High paid	Low paid	High paid
Age	36.71 (9.47)	31.28 (8.85)	38.11 (8.01)	31.99 (8.58)	38.36 (7.92)
Handicap	0.16	0.07	0.05	0.07	0.05
Married	0.51	0.37	0.70	0.44	0.70
Immigrant	0.39	0.31	0.19	0.31	0.19
Apprenticeship	0.43	0.55	0.46	0.60	0.45
Vocational training	0.15	0.10	0.24	0.11	0.24
University	0.09	0.11	0.21	0.09	0.21
Children	0.81	0.59	0.83	0.61	0.84
Local unemployment rate	10.10 (2.61)	9.64 (2.60)	9.43 (2.33)	9.64 (2.38)	9.42 (2.34)
Year of entry 2000	0.43	0.51	0.46	0.49	0.46
Number of observations	243	138	2585	255	2468

Source: SOEP, descriptives with respect to the year of entry, standard deviations in pharentheses, n=2,966
Threshold 1: 2/3 of the median, Threshold 2: First quintile

ployed in year t is around 15% and 13% and thus clearly higher for those who were low paid in $t-1$ than for previously high paid individuals with 2.4% and 2.2%, respectively. In addition to that, previously not employed individuals clearly have a higher probability of being low paid than previously high paid individuals. These descriptive statistics suggest that there may exist state dependence in all the three analyzed labor market states as well as a "low pay no pay cycle", independent of the threshold definition. These main results do not change fundamentally if attrition is additionally taken into account. However, not employed individuals leave the sample with the highest (8.6%) and high paid individuals with the lowest probability (5.4%), see Table 4.4. Attrition in this context means a drop out from the SOEP. The share of panel attrition is relatively low and underestimates the real panel attrition because the individuals have to be observed for two subsequent waves for entering the sample for estimation reasons, i.e. panel attrition in this context refers to drop out in the third or one of the following years of observation in my sample.

In Table 4.5 descriptive statistics of the observed characteristics are reported, conditioned on the labor market state in the first year of observation. Higher paid employees are on average better educated than non working and low paid individuals. The difference in education between non working and low paid individuals is relatively small. Moreover, low paid individuals are younger, are less often married and have fewer children than high paid and non working persons. The share of immigrants is higher among the low paid and not employed and the average local unemployment rate is higher among not employed individuals but the differences between the three groups are small. These results are quite similar for both thresholds.

The different aggregate transition probabilities for individuals in low and high paid jobs or in non-employment reported above could derive from observed and unobserved heterogeneity as well as from true state dependence, i.e. the fact that being low paid in one period itself increases the probability of being low paid in the next period (Stewart and Swaffield, 1999). If certain observable or unobservable individual characteristics go along with low transition probabilities into higher paid jobs, such as education or age, this will create aggregate state dependence although there does not need to be true state dependence. I will distinguish the different sources of the observed different transition probabilities in the econometric part of this paper and analyze whether and to which extent one can observe true state dependence in the three labour market states.

4.3 Modeling Transition Probabilities

This study analyzes the mobility between high pay ($j = 1$) and low pay employment ($j = 2$) on the one hand and non-employment ($j = 3$) on the other hand. Earnings are classified into two discrete ranges, low pay and high pay. I estimate the transition probabilities P between the three states from period $t-1$ to t. The transition matrix TM corresponds to

$$TM = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}. \tag{4.1}$$

I assume a first-order Markov process. The latent propensity E^* of individual i to be in state j in period t can be written as

$$E^*_{i,j,t} = X_{it}\beta_j + Z_{it-1}\gamma_j + \alpha_{ij} + \epsilon_{ijt}. \tag{4.2}$$

X_{it} contains individual observed characteristics in period t and Z_{it-1} contains the lagged labor market state, consisting of two dummy variables which indicate the state in period $t-1$ with high paid employment as the base category. Vector $\alpha_i = \{\alpha_{i1}, \alpha_{i2}, \alpha_{i3}\}$ describes the individual specific unobserved heterogeneity and ϵ_{ijt} is the error term. The error term is assumed to be independent from observable and unobservable individual characteristics and to follow a Type I extreme value distribution. The labor market state Z_{it} with the highest propensity $E^*_{i,j,t}$ is realized ($Z_{it} = j$ if $E^*_{i,j,t} > E^*_{i,l,t}$ for any $l \neq j$). This ends up in a multinomial logit panel data model with random effects with three states. Alternatively one could model the propensities to be employed and to be high or low paid simultaneously with two probit models ((Cappellari and Jenkins, 2004)). However, a disadvantage of this approach is that exclusion restrictions are required. Therefore, and because the three labor market states are mutually exclusive I choose a multinomial logit model. For other studies applying a multinomial logit model in the context of low pay dynamics see e.g. (Cappellari and Jenkins, 2004). Applying a standard multinomial logit model would imply the restrictive and often unrealistic assumption of independence of irrelevant alternatives (IIA), see e.g. Cameron and Trivedi (2005). With the introduction of random effects this assumption is relaxed as the random effects have to be integrated out and the denominators of the logit formula are inside the integral and therefore do not cancel out when calculating the probability ratio of two alternatives (Train, 2003). For a given unobserved heterogeneity the probability of individual i to be in state j in period t corresponds to

$$P(Z_{it} = j | X_{it}, Z_{it-1}, \alpha_i) = \frac{exp(X_{it}\beta_j + Z_{it-1}\gamma_j + \alpha_{ij})}{\Sigma_{k=1}^{3} exp(X_{it}\beta_k + Z_{it-1}\gamma_k + \alpha_{ik})}. \tag{4.3}$$

The coefficient vectors β_1 and γ_1 and the unobserved heterogeneity term α_{i1} of the base category are set to 0 for identification of the model.

The observation period of transition probabilities does not coincide with the start of the stochastic process generating individual's employment dynamics. Therefore, when modelling transition probabilities the initial condition problem has to be taken into account, see e.g. Heckman (1981a).

To take the problem of initial condition into account, I follow Gong, van Soest, and Villagomez (2004) and estimate a static multinomial logit model for the initial labor market state ($t = 0$) without lagged labor market states and different slope parameters similar to Heckman (1981b) estimating dynamic binary choice models.[5]

The probability of individual i to be in state j in the initial period $t = 0$ corresponds to

$$P(Z_{it} = j|X_{it}, \nu_i) = \frac{exp(X_{it}\delta_j + \nu_{ij})}{\Sigma_{k=1}^{J} exp(X_{it}\delta_k + \nu_{ik})} \tag{4.4}$$

with the unobserved heterogeneity $\nu_i = \{\nu_{i1}, \nu_{i2}, \nu_{i3}\}$ and the state specific coefficient vector δ_j. Being high paid ($j = 1$) is the base category and the coefficient vector δ_1 and the unobserved heterogeneity term ν_{i1} are set to 0.

The unobserved heterogeneity or random effects $\nu_i = \{\nu_{i2}, \nu_{i3}\}$ are functions of the unobserved heterogeneity α_i. Similar to Gong, van Soest, and Villagomez (2004) I assume that $\nu_i = C\alpha_i$, where C is a lower triangular matrix and has to be estimated.

$$C = \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix} \tag{4.5}$$

The unobserved heterogeneity $\alpha_i = \{\alpha_{i2}, \alpha_{i3}\}$ is unknown and has to be integrated out when calculating the likelihood function. In a model without

[5] For a discussion of the identification of multinomial discrete choice models with lagged dependent variables see (Honoré and Kyriazidou, 2000).

attrition the individual likelihood contribution can be written as

$$L_i = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \frac{exp(X_{it}\beta_2 + Z_{it-1}\gamma_2 + \alpha_2)^{l_t} exp(X_{it}\beta_3 + Z_{it-1}\gamma_3 + \alpha_3)^{n_t}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\beta_k + Z_{it-1}\gamma_k + \alpha_k)} \frac{exp(X_{it}\delta_2 + \nu_2)^{l_0} exp(X_{it}\delta_3 + \nu_3)^{n_0}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\delta_k + \nu_k)} f(\alpha) d\alpha \tag{4.6}$$

with $l_t = 1$ ($n_t = 1$) if the individual is low paid (not employed) in t, $l_t = 0$ ($n_t = 0$) if not and $l_0 = 1$ ($n_0 = 1$) if the individual is low paid (not employed) in the initial period and $l_0 = 0$ ($n_0 = 0$) if not.[6]

In general, panel attrition is not taken into account in studies dealing with employment dynamics. As long as the unobserved individual heterogeneity influencing the employment dynamics is not correlated with the unobserved term influencing the attrition process, no problem occurs. But a correlation of these terms could lead to biased estimates. In my data set, non-employment and low paid jobs go along with a higher probability of attrition, see Table 4.4. Therefore I take potential endogeneous sample attrition into account by estimating the employment transitions and the attrition process simultaneously. The latent attrition propensity D_t^* is assumed to be a linear function of the in the previous period observed characteristics A_{it-1} and unobserved characteristics α_{i4}. Attrition is present if the latent propensity D_t^* is positive.

$$D_t^* = A_{it-1}\xi_+ \alpha_{i4} + \kappa_{it} > 0 \tag{4.7}$$

The error terms κ_{it} are assumed to be independent from observed and unobserved characteristics and to follow a logistic distribution. This ends up in a logit model for the attrition equation. The indicator variable a_t takes on the value 1 if the individual is not interviewed in year t and 0 if no attrition occurs. For an individual with T observed years and the observation period ending before 2003 the corresponding likelihood contribution is given by

[6] An alternative estimator for dynamic discrete choice models is given by Wooldridge (2005) who propose to estimate the distribution conditional on the initial state and time invariant variables instead of jointly modeling all outcome variables. This ends up in less complex estimation methods. For an application in the context of dynamic multinomial discrete choice models see Haan (2005). However, for this approach a balanced panel is needed.

$$L_i = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \frac{exp(X_{it}\beta_2 + Z_{it-1}\gamma_2 + \alpha_2)^{l_t} exp(X_{it}\beta_3 + Z_{it-1}\gamma_3 + \alpha_3)^{n_t}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\beta_k + Z_{it-1}\gamma_k + \alpha_k)} \frac{exp(X_{it}\delta_2 + \nu_2)^{l_0} exp(X_{it}\delta_3 + \nu_3)^{n_0}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\delta_k + \nu_k)} \prod_{t=2}^{T+1} \Big(\frac{exp(A_{it-1}\xi + \alpha_4)^{a_t}}{1 + (A_{it-1}\xi + \alpha_4)} \Big) f(\alpha) d\alpha \tag{4.8}$$

For an individual with the last observation in 2003 no panel attrition occurs after entering into the sample. In this case the likelihood contribution can be written as

$$L_i = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \Big(\frac{exp(X_{it}\beta_2 + Z_{it-1}\gamma_2 + \alpha_2)^{l_t} exp(X_{it}\beta_3 + Z_{it-1}\gamma_3 + \alpha_3)^{n_t}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\beta_k + Z_{it-1}\gamma_k + \alpha_k)} \Big) \frac{exp(X_{it}\delta_2 + \nu_2)^{l_0} exp(X_{it}\delta_3 + \nu_3)^{n_0}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\delta_k + \nu_k)} \prod_{t=2}^{T} \Big(\frac{1}{1 + (A_{it-1}\xi + \alpha_4)} \Big) f(\alpha) d\alpha \tag{4.9}$$

For the estimation of the selection process into panel attrition, an identification restriction is needed. Therefore, A_{it-1} includes the same variables as X_{it-1} and Z_{it-1} and as the exclusion restriction a dummy variable indicating an interviewer change between $t-1$ and t. Interviewer changes are potentially endogenous with respect to wage mobility. For individuals moving due to a new job, one will probably observe an interviewer change. Therefore, I define an interviewer change only if the interviewer of the last year drops out of the SOEP interviewer sample, i.e. we do not observe any interviews of this interviewer in period t. An interviewer change defined in this way should be exogenous with respect to employment dynamics but should have a positive influence on the attrition probability. This influence should arise because the first meeting with an interviewer should go along with a relatively high tendency to refuse participation and subsequent contacts should increase trust. For a similar argument in the context of item nonresponse see Schräpler (2004).

I estimate a model with free correlations. The correlation coefficient ρ_1 measures the correlation between unobservable individual specific characteristics influencing the probability of being low paid and not employed in t while the correlations ρ_2 and ρ_3 describe the association between unobservables determining the attrition process and the probability of being low paid (ρ_2) and not employed (ρ_3). If $\rho_2 = \rho_3 = 0$, the attrition process can be assumed to be exogenous and the model reduces to a dynamic multinomal logit model as suggested by Gong, van Soest, and Villagomez (2004).

It is assumed that the individual specific random intercepts $\alpha_i = \{\alpha_{i2}, \alpha_{i3}, \alpha_{i4}\}$ follow a multivariate normal distribution. The likelihood contribution involves a 3 dimensional integration. I estimate the models with a Maximum Simulated Likelihood (MSL) approach. In this approach simulated probabilities are used instead of exact probabilities, see Gourieroux and Monfort (1993) or Hajivassiliou and Ruud (1994) for the properties of MSL.

In this MSL approach the integral in equation (4.8) is replaced by

$$L_i = \frac{1}{R}\sum_{d=1}^{R} \quad \prod_{t=1}^{T}\Big(\frac{exp(X_{it}\beta_2 + Z_{it-1}\gamma_2 + \alpha_2^d)^{l_t} exp(X_{it}\beta_3 + Z_{it-1}\gamma_3 + \alpha_3^d)^{n_t}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\beta_k + Z_{it-1}\gamma_k + \alpha_k^d)}\Big)$$
$$\frac{exp(X_{it}\delta_2 + \nu_{i2})^{l_0} exp(X_{it}\delta_3 + \nu_{i3})^{n_0}}{1 + \Sigma_{k=2}^{3} exp(X_{it}\delta_k + \nu_{ik})} \prod_{t=2}^{T+1}\Big(\frac{exp(A_{it-1}\xi_4 + \alpha_4^d)^{a_t}}{1 + (A_{it-1}\xi_4 + \alpha_4^d)}\Big)$$

For equation (4.9) the integral is replaced in the same way. In general independent random draws from mixing distributions are used in simulation approaches. In this paper I apply Halton Sequences as an alternative method, for details see e.g. Train (2003). The superior coverage compared to random draws and the negative correlation over the observations lead to a significant reduction in estimation time. For example Train (2000) and Bhat (2001) find in their studies that the results of mixed logit models are more precise with 100 Halton draws than with 1000 random draws. In this paper I use $r = 200$ Halton draws per individual. The models are programmed in Stata Version 8.2. A description of the applied simulation procedure in the context of random effects multinomial logit models is given in the Appendix, Chapter A.

Extent of State Dependence

The magnitudes of the estimated coefficients provide little information about the extent of true state dependence. State dependence describes the effect of being in one state compared to another state in $t-1$ on the probability of being in state j in period t. Therefore and due to the nonlinearity of the model, the measure of true state dependence SD is derived by calculating the average of pairwise individual differences between the predicted probabilities of being in state j conditional on two of the three labor market states. For example, the effect of being low paid ($j=2$) compared to being high paid ($j=1$) in $t-1$ on the probability of being low paid in t can be written as

$$SD = \frac{1}{N}\sum_{i=1}^{N}(P_i(j=2|j=2) - P_i(j=2|j=1)). \tag{4.10}$$

In order to derive the individual specific probabilities for each category given observed and unobserved characteristics it is necessary to assign individual values to the random intercepts. An individual value is given by the mean of the individual specific posterior distribution of unobserved heterogeneity. The posterior distribution depends on the prior (estimated) distribution of unobserved heterogeneity and the observed individual sequence of labor market states. This way of assigning values to latent variables is sometimes referred to as *Empirical Bayes* prediction (Skrondal and Rabe-Hesketh, 2004).[7]

I follow Train (2003) and take r draws of α from the population distribution and calculate the individually weighted averages of these draws. The weight for each draw d is proportional to the probability of the observed sequence of labor market states $P(y_i|x_i, \alpha_d)$. The simulated individual mean $\tilde{\alpha}_i$ is given by:

$$\tilde{\alpha}_i = \sum_{d=1}^{R} w^d \alpha^d \tag{4.11}$$

The higher the probability of the chosen sequence given the unobserved

[7]Alternatively one could use the expected value 0 of the unobserved heterogeneity for all individuals. However, in this study the extent of state dependence of different groups, e.g. the initially low paid individuals in my sample, is of interest and the average latent values probably vary with the initial state which may have a relevant influence on the predictions.

heterogeneity the higher the weight w^d assigned to the draw:

$$w^d = \frac{P(y_i|x_i, \alpha_d)}{\sum_{d=1}^{R} P(y_i|x_i, \alpha_d)}$$

Given the unobserved and observed heterogeneity, individual transition probabilities between the three states can be calculated. The standard errors of average transition probabilities and of extents of true state dependence are computed using parametric bootstrap.

4.4 Results

I estimate dynamic multinomial logit panel data models with random effects and potential endogenous panel attrition for two different low pay thresholds. The results of the dynamic equations and the distributions of the unobserved heterogeneity are reported in Tables 4.6 and 4.7, the results of the static multinomial logit model and the attrition process are reported in the Appendix, Tables 4.11 and 4.12. For both thresholds I estimate the process with (model 2) and without unobserved heterogeneity (model 1). In the following I compare the different models and evaluate the endogeneity of the initial state and the attrition process. Second, I report the coefficients of the models and third I discuss the extent of true state dependence.

4.4.1 Endogeneity of Initial State and Attrition

Both correlation coefficients describing the unobserved heterogeneity of the attrition process and the probability of being low paid and the probability of non-employment, respectively, are not significantly different from 0. This indicates that panel attrition is exogenous with respect to low pay and non-employment dynamics and the employment dynamics and the attrition process can be estimated separately. According to that, the results of a dynamic multinomial logit model without simultanous estimation of the attrition process are very similar to the one of the full model, see Table 4.13 in the Appendix.

Compared to a simple multinomial logit model (model 1) the inclusion of unobserved heterogeneity and the modeling of the initial condition (model 2) significantly increase the log-likelihood and clearly reduces the coefficients of the lagged labor market state variables. These results confirm previous research on low pay and unemployment dynamics and emphasize the importance of the initial condition problem within dynamic panel data models. For both thresholds the correlation coefficient ρ_{12} is around 0.7, indicating that unobserved characteristics which lead to low pay employment and non-employment are similar but different from unobserved characteristics of high paid individuals. The estimated variances of all random intercepts are significant and the point estimates of the variances in the dynamic equation (4.81 and 8.37 for the threshold 1 and 4.95 and 8.14 for the threshold 2) indicate that both random intercepts contribute more to the state probability than the idiosyncratic errors with a normalized variance of $\pi^2/6$.

Table 4.6: Dynamic Multinomial Logit Model, Threshold 1, joint estimation with the Attrition Process

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1				Model 2			
	Low Paid		Non-employment		Low Paid		Non-employment	
Low paid, t-1	3.37**	0.16	2.23**	0.18	1.17**	0.29	0.93**	0.35
Non Employment, t-1	3.14**	0.18	4.73**	0.13	1.49**	0.30	2.31**	0.28
Year 2000	-0.12	0.23	0.20	0.21	-0.05	0.26	0.13	0.24
Year 2001	0.08	0.20	0.39*	0.18	0.13	0.23	0.34	0.23
Year 2002	0.52**	0.20	0.71**	0.18	0.61**	0.23	0.73**	0.23
Year 2003	0.07	0.22	0.80**	0.18	0.24	0.25	0.96**	0.24
Age	-0.18**	0.07	-0.15*	0.06	-0.49**	0.11	-0.43**	0.11
Age squared $*10^{-2}$	0.22*	0.09	0.22**	0.08	0.58**	0.14	0.57**	0.14
Apprenticeship	-0.27	0.17	-0.73**	0.14	-0.79**	0.28	-1.84**	0.32
Vocational training	-0.90**	0.24	-1.05**	0.18	-1.68**	0.36	-2.51**	0.39
University	-1.23**	0.27	-1.54**	0.21	-2.29**	0.41	-3.39**	0.47
Non German	0.58	0.15	0.67**	0.13	1.12**	0.24	1.57**	0.28
Married	-0.53**	0.16	-0.70**	0.14	-0.90**	0.23	-1.34**	0.24
Children	0.02	0.08	0.16*	0.06	0.10	0.10	0.27*	0.10
Handicap	0.43	0.24	1.09**	0.16	1.03**	0.33	2.12**	0.31
Local unemp. rate	0.12**	0.02	0.06**	0.02	0.21**	0.04	0.17**	0.04
Constant	-0.98	1.25	-1.58	1.17	3.94*	1.94	2.31	2.11
						Coef.	Std. Err.	
σ_1^2		-	-			4.81**	1.20	
σ_2^2		-	-			8.37**	2.03	
σ_3^2		-	-			3.50**	1.29	
ρ_{12}		-	-			0.70**	0.09	
ρ_{23}		-	-			-0.01	0.10	
ρ_{33}		-	-			-0.12	0.15	
Log-Likelihood		-5,668.57				-5,408.37		

The unobserved heterogeneity is assumed to follow a multivariate normal distribution. The equations of the initial state and the attrition process are reported in the Appendix. Observations: 2,966 individuals.

Model 1: No unobserved heterogeneity; Model 2: Jointly distributed unobserved heterogeneity.

Threshold 1: 2/3 of the median wage; Threshold 2: first quintile of the wage distribution.

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Table 4.7: Dynamic Multinomial Logit Model, Threshold 2, joint estimation with the Attrition Process

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1				Model 2			
	Low Paid		Non-employment		Low Paid		Non-employment	
Low paid, t-1	3.38**	0.13	2.35**	0.16	1.25**	0.24	1.11**	0.30
Non Employment, t-1	3.34**	0.17	5.02**	0.14	1.76**	0.28	2.62**	0.29
Year 2000	0.07	0.19	0.27	0.21	0.05	0.22	0.19	0.25
Year 2001	0.14	0.16	0.41*	0.18	0.13	0.20	0.38	0.23
Year 2002	0.43*	0.17	0.76**	0.18	0.33	0.20	0.76**	0.24
Year 2003	-0.02	0.19	0.84**	0.18	-0.08	0.22	0.98**	0.24
Age	-0.06	0.06	-0.12	0.06	-0.34**	0.10	-0.39**	0.11
Age squared $*10^{-2}$	0.07	0.08	0.18*	0.08	0.39**	0.12	0.52**	0.14
Apprenticeship	-0.26	0.15	-0.75**	0.14	-0.77**	0.25	-1.85**	0.32
Vocational training	-0.96**	0.21	-1.08**	0.19	-1.79*	0.32	-2.53**	0.40
University	-1.18**	0.23	-1.58**	0.21	-2.39**	0.38	-3.42**	0.47
Non German	0.51**	0.13	0.66**	0.13	1.13**	0.22	1.57**	0.27
Married	-0.60**	0.14	-0.72**	0.14	-1.00**	0.20	-1.30**	0.24
Children	0.01	0.06	0.15*	0.06	0.08	0.09	0.22*	0.10
Handicap	0.35	0.21	1.10**	0.17	0.85**	0.30	2.12**	0.31
Local unemp. rate	0.08**	0.02	0.06*	0.02	0.16**	0.04	0.16**	0.04
Constant	-2.79*	1.10	-2.30	1.18	2.23	1.78	1.60	2.11
						Coef.	Std. Err.	
σ_1^2		-	-			4.95**	1.11	
σ_2^2		-	-			8.14**	2.01	
σ_3^2		-	-			3.31**	1.32	
ρ_{12}		-	-			0.68**	0.08	
ρ_{23}		-	-			-0.12	0.10	
ρ_{33}		-	-			-0.05	0.18	
Log-Likelihood		-6,206.31				-5,933.66		

The unobserved heterogeneity is assumed to follow a multivariate normal distribution. The equations of the initial state and the attrition process are reported in the Appendix. Observations: 2,966 individuals

Model 1: No unobserved heterogeneity; Model 2: Jointly distributed unobserved heterogeneity

Threshold 1: 2/3 of the median wage; Threshold 2: first quintile of the wage distribution.

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

4.4.2 Model Estimates

Several covariables are included in the regressions. The results indicate that immigrants have a higher probability of being low paid or not employed, while married men are more often in high paid jobs. The existence of children in the household goes along with a higher probability of being not employed, while the coefficient of having children is not significantly different from zero with respect to the probability of being low-paid. The age has a U-shaped influence on the probability of being low paid and not employed: The younger and the older persons have a lower probability of being in a higher paid job. A higher education goes along with a higher probability of being in a high paid job. The comparison group of the three categories “apprenticeship”, “further vocational training” and “university” consists of individuals with no vocational training at all. Moreover, individuals with a handicap have a higher risk of non-employment or being low paid and a higher local unemployment rate decreases the probability of being high paid.

The coefficients of the lagged labor market states indicate that there exists true state dependence in low pay as well as in non-employment in west Germany for men for both thresholds (see Tables 4.6 and 4.7). Being low paid in year $t-1$ increases the probability of being low paid compared to the probability of being high paid in year t. Being not employed in year $t-1$ increases the probability of being not employed compared to the probability of being high paid in year t.

In addition to that there exists a strong relation between low pay and no pay. Being low paid in year $t-1$ increases the probability of non working compared to the probability of being high paid in year t. Being not employed in year $t-1$ increases the probability of being low paid compared to the probability of being high paid.

4.4.3 Extent of True State Dependence

As mentioned above, the coefficients provide little information about the extent of true state dependence. Table 4.8 contains the transition matrices between the three states for both thresholds, based on averaged transition probabilities across all individuals. Independent of the previous labor market state, the average probability to be high paid in the next period is above 80%. This result holds for both thresholds and can be explained by the influence of observable and unobservable characteristics shifting the main share of individuals into relatively high paid jobs, independent of their employment state in the last year. However, the probability of being high paid is with 93% and 92%, respectively, the highest for individuals who have been high paid in the previous period. Previous non-employment goes along with a probability of 84% (80%) and previous low payment with a probability of 89% (86%) to be high paid in t.

Table 4.8: Estimated Transition Matrix: all Men

All men	Non-employment, t	Low paid, t	High paid, t
	Threshold 1		
Non-employment, t-1	10.98	4.80	84.22
	(9.41-13.52)	(3.43-6.42)	(81.21-86.33)
Low paid, t-1	6.26	4.95	88.79
	(4.55-7.74)	(3.62-6.74)	(86.38-90.70)
High paid, t-1	4.36	2.43	93.22
	(3.46-5.25)	(1.77-3.15)	(92.21-94.30)
	Threshold 2		
Non-employment, t-1	11.53	8.39	80.08
	(9.48-14.98)	(6.42-10.82)	(76.16-82.98)
Low paid, t-1	6.21	7.74	86.05
	(4.74-7.34)	(6.24-9.75)	(83.99-87.86)
High paid, t-1	4.09	3.80	92.11
	(3.14-5.12)	(2.93-4.71)	(90.87-93.34)

Source: SOEP, waves 1998-2003, n=2,966
Threshold 1: 2/3 of the median, Threshold 2: First quintile
The 5th and 95th percentiles are given in parentheses, derived using parametric bootstrap with 200 replications.

Table 4.9 contains the transition probabilities for three groups defined by their initially observed state. Compared to Table 4.8 the results change and are similar to the descriptive transition matrices. For example more than 65% of the sample consisting of initially not employed individuals are not employed in the subsequent period, conditional on non-employment in the previous period, and the predicted probability of staying low paid is around 40% for the group of initially low paid individuals.

The differences in the state probabilities can be attributed to the different previous labor market states and therefore provide information about true state dependence. Because the extent of state dependence may differ with respect to observed and unobserved heterogeneity, I calculated the SD of four different groups separately: all men, initially not employed, low paid and high paid men. The results are presented in Table 4.10.

For the whole sample being not employed increases the probability of being not employed in the future by 6.63% for threshold 1 and 7.44% for threshold 2. This state dependence is higher compared to the state dependence in low paid jobs (2.52% and 3.94%, respectively). Moreover, low paid jobs increase the probability to be not employed by 1.90% and 2.12%, while not employed

Table 4.9: Estimated Transition Matrices: selected samples with respect to the initial state

	Non-employment, t	Low paid, t	High paid, t
Initially not employed		Threshold 1	
Non-employment, t-1	67.67	11.22	21.11
	(64.00-71.23)	(8.71-14.13)	(18.16-23.48)
Low paid, t-1	47.94	16.28	35.79
	(34.87-56.93)	(10.49-23.46)	(28.34-45.54)
High paid, t-1	37.71	9.39	52.91
	(28.08-46.44)	(6.03-13.46)	(44.44-63.24)
		Threshold 2	
Non-employment, t-1	67.20	15.37	17.43
	(63.33-70.76)	(12.64-18.79)	(14.71-19.68)
Low paid, t-1	46.74	20.70	32.56
	(36.45-54.72)	(14.44-27.95)	(26.15-41.44)
High paid, t-1	35.21	12.17	52.62
	(26.17-44.95)	(8.09-16.71)	(43.29-62.62)
Initially low paid		Threshold 1	
Non-employment, t-1	21.84	40.11	38.05
	(14.74-31.04)	(30.73-50.31)	(28.85-46.27)
Low paid, t-1	10.25	40.83	48.92
	(7.91-12.63)	(35.57-46.58)	(42.67-54.72)
High paid, t-1	7.40	22.84	69.76
	(4.69-10.33)	(15.80-31.24)	(61.01-77.46)
		Threshold 2	
Non-employment, t-1	18.51	45.96	35.53
	(13.10-25.48)	(37.41-54.88)	(26.93-43.05)
Low paid, t-1	8.57	43.08	48.35
	(6.74-10.30)	(38.69-47.69)	(43.43-53.03)
High paid, t-1	5.75	23.45	70.80
	(3.72-7.82)	(17.57-30.58)	(63.81-76.89)
Initially high paid		Threshold 1	
Non-employment, t-1	5.07	2.30	92.63
	(3.63-7.82)	(1.28-3.71)	(89.59-94.73)
Low paid, t-1	2.12	1.96	95.91
	(1.30-3.19)	(1.04-3.31)	(93.97-97.26)
High paid, t-1	1.05	0.68	98.26
	(0.83-1.26)	(0.52-0.88)	(97.96-98.53)
		Threshold 2	
Non-employment, t-1	5.32	3.81	90.87
	(3.46-8.77)	(2.33-5.76)	(86.97-93.55)
Low paid, t-1	1.97	2.81	95.23
	(1.27-2.88)	(1.78-4.38)	(93.34-96.49)
High paid, t-1	0.85	0.95	98.20
	(0.66-1.02)	(0.77-1.16)	(97.90-98.46)

Source: SOEP, waves 1998-2003, n=2,966

Threshold 1: 2/3 of the median, Threshold 2: First quintile

The 5th and 95th percentiles are given in parentheses, derived using parametric bootstrap with 200 replications. Transition probabilities are calculated separately for three groups, defined by their initially observed state.

Table 4.10: Estimated State Dependence (SD)

	Threshold 1			
	All men	Not employed (t_0)	Low paid (t_0)	High paid (t_0)
SD LP	2.52	6.89	17.99	1.28
	(0.97-4.23)	(1.74-12.12)	(9.23-26.37)	(0.44-2.49)
SD NP	6.63	29.96	14.43	4.02
	(4.54-9.97)	(21.51-40.16)	(8.97-21.89)	(2.56-6.81)
SD NP-LP	2.37	1.83	17.28	1.62
	(0.78-4.32)	(-2.54-6.00)	(7.60-27.51)	(0.69-2.96)
SD LP-NP	1.90	10.23	2.85	1.07
	(0.10-3.76)	(-0.59-20.86)	(-0.70-6.10)	(0.26-2.09)
	Threshold 2			
	All men	Not employed (t_0)	Low paid (t_0)	High paid (t_0)
SD LP	3.94	8.52	19.63	1.86
	(2.06-6.01)	(3.28-13.71)	(12.08-27.04)	(0.85-3.27)
SD NP	7.44	31.99	12.76	4.47
	(4.82-11.65)	(22.07-42.84)	(13.23-32.40)	(2.56-7.93)
SD NP-LP	4.59	3.19	22.51	2.87
	(2.21-7.33)	(-1.72-8.10)	(13.23-32.40)	(1.48-4.74)
SD LP-NP	2.12	11.53	2.82	1.12
	(0.50-3.70)	(2.10-20.42)	(0.29-5.27)	(0.41-1.94)

Source: SOEP, waves 1998-2003, n=2,966
Threshold 1: 2/3 of the median, Threshold 2: First quintile
The 5th and 95th percentiles are given in parentheses, derived using parametric bootstrap with 200 replications.
SD: State Dependence; LP: Low Pay; NP: No Pay;

individuals have a higher probability to be low paid in the next period (2.37% and 4.59%, respectively). There is evidence for a "low pay no pay cycle", but the individuals seem to be better off if they have a low paid job than no job at all. This can also be seen in Table 4.8, indicating that low pay employment leads with a significantly higher probability (88.79%) to higher paid jobs than non-employment (84.22%).

The extent of the state dependence varies between the groups. The initially high paid men are characterized by the lowest marginal effects of the lagged states, while initially low paid and not employed men have relatively strong effects of state dependence in low pay and non-employment. For example the extent of state dependence in low paid jobs (SD LP) for initially not employed men is 6.89% (8.52%) and the corresponding effect in non-employment is 29.96% (12.76%).

Although there exists evidence for a "low pay - no pay cycle", the estimated effect of previous non-employment on the probability of non-employment is significantly higher for all groups and both thresholds than the effect of being previously low paid, see Tables 4.8 and 4.9. Moreover the point estimates to be high paid are always higher for previous low payment, although this difference is not always significantly different from zero. However, for the whole sample the confidence bands do not overlap, see Table 4.8.

I find some evidence that low paid jobs are stepping stones to better jobs in west Germany and the results indicate that being low paid does not have any adverse effects on future employment prospects if it is compared with non-employment. Thus, these results are not consistent with the hypothesis that a low-wage job does not augment a person's human capital more than unemployment. The results allow a more positive evaluation of low wage employment than the results of Stewart (2006) who does not estimate different effects of previous unemployment and previous low paid jobs on the probability of unemployment. But in comparison to high paid jobs being low paid goes along with a higher risk of non-employment and a higher probability of being low paid in the future.

4.5 Conclusion

This paper examines the low pay and non-employment dynamics of men in west Germany. I estimate a dynamic multinomial logit model with random effects and take the the initial condition problem into account. In addition to that I take potential endogeneity of panel attrition into account by estimating the processes of panel attrition and employment dynamics simultaneously. There is no evidence of endogeneous panel attrition, the corresponding correlation coefficients are insignificant and the results do not change compared to the simpler model.

This first study on low pay dynamics in Germany indicates that there exists strong true state dependence in low pay as well as in non-employment for men in west Germany. In addition to that there is a strong link between low pay and no pay. Despite this evidence for a "low pay no pay cycle", compared to non-employment low-wage jobs increase the probability of being employed in the future. Moreover, low paid jobs seem to lead to a higher paid job in the future.

This study finds some evidence that low paid jobs are stepping stones to better jobs in west Germany and no evidence that being low paid does have any adverse effects on future employment prospects if it is compared with non-employment. However, in comparison to high paid jobs being low paid goes along with a higher risk of non-employment and a higher probability of being low paid in the future.

4.6 Appendix

Table 4.11: Dynamic Multinomial Logit Model, Threshold 1: Initial State equation and Attrition Process

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1						Model 2					
	Initial State				Attrition		Initial State				Attrition	
	Low paid		Non-employment				Low paid		Non-employment			
Low paid, t-1					0.20	0.21	-	-	-	-	0.24	0.37
Non-employment, t-1					0.44**	0.15	-	-	-	-	0.94	0.51
Year 2000	0.51**	0.19	0.26	0.15	-	-	0.52	0.29	0.03	0.26	-	-
Year 2001	-	-	-	-	0.01	0.15	-	-	-	-	0.39	0.21
Year 2002	-	-	-	-	1.00**	0.13	-	-	-	-	0.60**	0.20
Year 2003	-	-	-	-	0.69**	0.14	-	-	-	-	0.57*	0.28
Age	-0.43**	0.09	-0.23**	0.07	-0.08	0.05	-0.81**	0.16	-0.60**	0.15	-0.18*	0.08
Age squared $*10^{-2}$	0.49**	0.12	0.30**	0.09	0.09	0.07	0.93**	0.21	0.74**	0.20	0.20*	0.10
Apprenticeship	-0.66**	0.24	-1.07**	0.19	-0.14	0.14	-1.37**	0.42	-2.02**	0.41	-0.21	0.23
Vocational training	-1.34**	0.35	-1.39**	0.23	-0.04	0.16	-2.73**	0.58	-3.06**	0.53	-0.04	0.25
University	-1.09**	0.34	-1.74**	0.27	-0.01	0.16	-2.30**	0.59	-3.21**	0.57	0.06	0.26
Non German	0.75**	0.22	0.90**	0.17	0.23*	0.11	1.52**	0.37	1.84**	0.36	0.34*	0.17
Married	-0.93**	0.24	-1.04**	0.18	-0.12	0.11	-1.46**	0.37	-1.63**	0.34	-0.16	0.16
Children	0.03	0.12	0.16*	0.08	-0.13*	0.05	0.09	0.16	0.24	0.14	-0.19*	0.08
Handicap	0.82*	0.38	1.42**	0.22	-0.12	0.19	1.64**	0.57	2.41**	0.48	-0.29	0.28
Local unemp. rate	0.09*	0.04	0.13**	0.03	-0.03	0.02	0.18**	0.06	0.24**	0.06	-0.01	0.03
Interviewer Drop out	-	-	-	-	0.68**	0.16	-	-	-	-	0.86**	0.22
Constant	5.65**	1.54	1.73	1.32	-1.27	0.97	10.82**	2.65	6.23*	2.64	0.13	1.50
		Coef.	Std. Err.					Coef.	Std. Err.			
ca		-	-					1.43**	0.29			
cb		-	-					0.65*	0.27			
cc		-	-					0.82**	0.14			

Unobserved heterogeneity is assumed to follow a multivariate normal distribution. Observations: 2,966 individuals.

Model 1: No unobserved heterogeneity; Model 2: Jointly distributed unobserved heterogeneity.

Threshold 1: 2/3 of the median wage; Threshold 2: first quintile of the wage distribution.

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Table 4.12: Dynamic Multinomial Logit Model, Threshold 2: Initial State equation and Attrition Process

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Model 1						Model 2					
	Initial State				Attrition		Initial State				Attrition	
	Low paid		Non-employment				Low paid		Non-employment			
Low paid, t-1	-	-	-	-	0.32	0.17	-	-	-	-	0.63	0.33
Non-employment, t-1	-	-	-	-	0.46**	0.15	-	-	-	-	0.78	0.55
Year 2000	0.43**	0.15	0.28	0.15	-	-	0.41	0.23	0.09	0.26	-	-
Year 2001	-	-	-	-	0.00	0.15	-	-	-	-	0.37	0.21
Year 2002	-	-	-	-	1.01**	0.13	-	-	-	-	0.59**	0.21
Year 2003	-	-	-	-	0.69**	0.14	-	-	-	-	0.55	0.28
Age	-0.34**	0.07	-0.26**	0.07	-0.08	0.05	-0.67**	0.13	-0.61**	0.15	-0.16	0.08
Age squared $*10^{-2}$	0.36**	0.10	0.32**	0.10	0.00	0.00	0.73**	0.17	0.74**	0.19	0.00	0.00
Apprenticeship	-0.37	0.20	-1.06**	0.19	-0.14	0.14	-0.95**	0.35	-1.91**	0.39	-0.21	0.23
Vocational training	-1.18**	0.26	-1.43**	0.23	-0.03	0.16	-2.34**	0.45	-3.03**	0.50	-0.03	0.25
University	-1.19**	0.28	-1.79**	0.27	0.00	0.16	-2.34**	0.48	-3.27**	0.54	0.08	0.26
Non German	0.79**	0.17	0.96**	0.17	0.22*	0.11	1.57	0.31	1.94**	0.35	0.33	0.17
Married	-0.61**	0.18	-1.05**	0.18	-0.12	0.11	-1.01**	0.28	-1.50**	0.31	-0.14	0.16
Children	-0.07	0.09	0.15	0.08	-0.13*	0.05	-0.07	0.13	0.15	0.13	-0.19*	0.08
Handicap	0.82**	0.29	1.47**	0.22	-0.13	0.19	1.68**	0.46	2.57**	0.46	-0.24	0.28
Local unemp. rate	0.09**	0.03	0.13**	0.03	-0.03	0.02	0.17**	0.05	0.25**	0.06	-0.01	0.03
Interviewer Drop out	-	-	-	-	0.69**	0.16	-	-	-	-	0.85**	0.22
Constant	4.77**	1.25	2.27	1.33	-1.41	0.97	9.80**	2.24	6.89**	2.58	-0.13	1.47
		Coef.	Std. Err.					Coef.	Std. Err.			
ca		-	-					1.31**	0.22			
cb		-	-					0.61**	0.23			
cc		-	-					0.84**	0.14			

Unobserved heterogeneity is assumed to follow a multivariate normal distribution. Observations: 2,966 individuals.

Model 1: No unobserved heterogeneity; Model 2: Jointly distributed unobserved heterogeneity.

Threshold 1: 2/3 of the median wage; Threshold 2: first quintile of the wage distribution.

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Table 4.13: Dynamic Multinomial Logit Model, Estimation with unobserved heterogeneity and without the attrition process

	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	Threshold 1				Threshold 2			
	Low paid		Non-employment		Low paid		Non-employment	
Low paid, t-1	1.10**	0.29	1.04**	0.35	1.23**	0.24	1.11**	0.30
Non-employment, t-1	1.67**	0.33	2.41**	0.28	1.80**	0.28	2.60**	0.29
Year 2000	-0.03	0.26	0.16	0.24	0.06	0.22	0.20	0.25
Year 2001	0.15	0.23	0.38	0.22	0.16	0.20	0.40	0.22
Year 2002	0.64**	0.23	0.78**	0.22	0.37	0.20	0.78**	0.23
Year 2003	0.26	0.25	1.02**	0.23	-0.04	0.22	1.00**	0.23
Age	-0.50**	0.11	-0.41**	0.11	-0.34**	0.10	-0.40**	0.11
Age squared $*10^{-2}$	0.59**	0.14	0.55**	0.14	0.39**	0.14	0.55**	0.14
Apprenticeship	-0.73**	0.27	-1.83**	0.32	-0.78**	0.25	-1.88**	0.32
Vocational training	-1.59**	0.35	-2.46**	0.38	-1.80**	0.32	-2.56**	0.40
University	-2.22**	0.42	-3.34**	0.46	-2.41**	0.38	-3.51**	0.48
Non German	1.09**	0.24	1.54**	0.27	1.14**	0.22	1.61**	0.27
Married	-0.89**	0.23	-1.30**	0.24	-0.99**	0.20	-1.34**	0.24
Children	0.12	0.10	0.23*	0.10	0.08	0.09	0.23*	0.10
Handicap	0.95**	0.34	2.08**	0.30	0.84**	0.30	2.14**	0.31
Local unemp. rate	0.20**	0.04	0.16**	0.04	0.15**	0.04	0.16**	0.04
Constant	3.98*	1.97	2.04	2.10	2.24	1.78	1.75	2.11
		Coef.	Std. Err.			Coef.	Std. Err.	
σ_1^2		4.69**	1.14			4.90**	1.07	
σ_2^2		7.61**	1.79			8.26**	1.96	
ρ_{12}		0.60**	0.11			0.66**	0.08	
Log-Likelihood	-3,442.95				-3,969.36			

Unobserved heterogeneity is assumed to follow a multivariate normal distribution. Observations: 2,966 individuals.

The estimation results for the initial state are not reported here and are quite similar to the one of the joint estimation with the attrition process.

Threshold 1: 2/3 of the median wage; Threshold 2: first quintile of the wage distribution.

*: statistically significant at least at the 5% level; **: statistically significant at least at the 1% level.

Chapter 5

Conclusion

Principal findings and policy conclusions

Reducing the high unemployment rate is one of the most important and challenging issues facing the German society. The risk of unemployment is especially high among low-skilled and unskilled individuals and higher among migrants than among natives. This book contributes to the ongoing debate about the determinants of individual employment dynamics by analyzing transitions processes between employment, unemployment and transfer receipt on the German labor market. The empirical studies are based on the German Socio Economic Panel Study (SOEP) applying duration and Markov models.

In Chapter 2, I estimate the influence of the ratio between potential labor income and the welfare payment level on the probability of moving from social assistance to work. The potential net labor income is estimated with standard wage equations by accounting for sample selection and applying a simple tax function. I estimate a discrete-time proportional hazard rate model with competing risks and risk specific unobserved heterogeneity. The ratio between potential labor income and welfare level shows a positive effect on the probability of moving to employment for households whose potential labor income exceeds their welfare payment level. This result is contrary to previous studies dealing with the determinants of social assistance spell duration in Germany. It derives from a simultaneous consideration of both sources of income, the net household labor income and the social welfare level, and additionally from a differentiation between transitions to work and alternative transitions. A reduction of the benefit level could be one solution to overcome the incentive problems. However, the amount of social assistance is related to a basic minimum income concept and a general reduction of benefit levels seems to be an unlikely political option. A reduction of social assistance levels is not the only way to overcome incentive problems of a transfer program, there exist other possible solutions like workfare.

There is much concern about the very high unemployment rates among migrants. Chapter 3 investigates the differences in unemployment dynamics between natives and migrants in Germany to provide evidence about the most relevant factors. The analysis is based on an inflow sample of men into unemployment and the estimation of a bivariate hazard rate model with two states, unemployment and employment. Two processes are analyzed: transitions from unemployment to employment and transitions from employment to unemployment. The durations of both states are estimated jointly and state specific unobserved heterogeneity components are allowed to be correlated across the two states. This is important because there is no reason to believe that unobserved characteristics determining the duration of unemployment are independent from unobserved characteristics influencing subsequent employment stability. Ignoring potential dependence could create a sample selection problem and thereby produce biased estimates. I find some evidence that the two processes are not independent from each other, but the results do not change qualitatively compared to a model with uncorrelated unobserved heterogeneity.

The results show that migrants stay longer unemployed than natives, and the probability of leaving unemployment differs strongly with ethnicity. While immigrants from Italy, Ex-Yugoslavia and Spain do not differ from natives, Turkish immigrants have a significantly lower probability of leaving unemployment for a paid job. Moreover, Turkish members of the second generation of guest-workers have a significantly lower probability of leaving unemployment than natives as well. However, once migrants find a new job, we observe no significant differences in their employment stability compared to natives, independent of ethnicity. These results suggest that, compared to natives with the same observable and unobservable characteristics, unemployed immigrants do not find less stable jobs but they need more time to find these jobs. Predominantly Turkish migrants from the first and second generation face the problem of slow integration from unemployment to employment. Therefore, adequate policy measures should concentrate on the job finding process of Turkish migrants to decrease their disadvantages on the labor market.

Chapter 4 examines the low pay and non-employment dynamics of men in west Germany. I estimate a dynamic multinomial logit model with random effects accounting for initial conditions. In addition to that, I consider potential endogeneity of panel attrition by estimating the processes of panel attrition and employment dynamics simultaneously. There is no evidence of endogeneous panel attrition, the corresponding correlation coefficients are insignificant and the results do not change compared to the simpler model. This first study on low pay dynamics in Germany indicates that there exists strong true state dependence in low pay as well as in non-employment for men in west Germany. Moreover, there is a strong link between low pay and no pay.

Despite this evidence for a “low pay - no pay” cycle, compared to non-employment low-wage jobs increase the probability of being employed in the future. Moreover, low paid jobs seem to lead to a higher paid job in the future. This study finds some evidence that low paid jobs are stepping stones to better jobs in west Germany and no evidence that being low paid has any adverse effects on future employment prospects if it is compared with non-employment. The results allow for a more positive evaluation of low wage employment than the results of Stewart (2006) who does not estimate different effects of previous unemployment and previous low paid jobs on the probability of unemployment for the UK. However, my results also show that in comparison to high paid jobs, being low paid goes along with a higher risk of non-employment and a higher probability of being low paid in the future.

Further research

This study is part of a large and still growing literature of empirical research on individual labor market transitions. Moreover, policy changes, in the context of employment dynamics in the low wage sector, for example the introduction of a new form of unemployment and social assistance (*Arbeitslosengeld II*), lead to new research questions. This implies that there are numerous ways and lines along which to extend and to improve the analyses presented here.

The analysis on duration of social assistance receipt gives interesting insights concerning the determinants of welfare dependence. However, the study cannot explain why households enter social assistance. One possibility to analyze the process of entering social assistance is the joint estimation of the duration of welfare receipt and the probability of reentering social assistance after leaving for a job, similar to the bivariate hazard rate model applied in Chapter 3. However, the sample size in the SOEP is too small to facilitate such an approach. Therefore, a better strategy is to model a Markov process similarly to the methods applied in Chapter 4 to analyze the probability of entering and leaving social assistance jointly. In addition to that, fully dynamic structural models following Keane and Wolpin (2002) could be a further extension. Moreover, from a political perspective, it would be interesting to analyze the effects of the policy changes in the transfer system introduced by the so called Hartz-Reforms in 2005.

My analysis of unemployment dynamics of migrants shows that Turkish migrants stay longer unemployed than natives and other migrants. In this context one question naturally arises: why do Turkish migrants stay longer unemployed after controlling for observed and unobserved characteristics than natives and other migrants? To answer this question, one needs more information about the search behavior of migrants and natives. For the UK, Frijters, Shields, and Price (2005) make use of such information. The SOEP contains questions about the search behavior since the wave 2003, i.e. similar research will be possible for Germany in the near future as well. In addition, there exist evidence that migrants have a lower probability to take part in labor market training programs than natives. Recent studies indicate a positive effect at least for some training programs (Schneider and Uhlendorff, 2006). Therefore, why and in how far migrants are treated differently and whether this has an effect on employment prospects are important research questions.

Chapter 4 includes methodological extensions of previous approaches in the field of low pay dynamics. However, additional methodological extensions are possible. They range from more flexible modeling of the transition processes by allowing for a higher order Markov process, or allowing for autocorrelation in the error terms as further potential source of state dependence to a different specification of the unobserved heterogeneity with less parametric assumptions. Beyond these methodological extensions, there exist interesting further research questions. For example, differences in the low pay dynamics between east and west Germany, whether or not state dependence has changed compared to the 1980s and whether low pay dynamics differs between men and women. In addition, the reform of the social and unemployment assistance (*Arbeitslosengeld II*) goes along with different incentive schemes in the low wage sector. It is important to know whether the presented results hold under this new regime and being low paid still improves future employment prospects if it is compared to non-employment.

Appendix A

Estimation of Multinomial Logit Models with unobserved Heterogeneity using Maximum Simulated Likelihood[1]

A.1 Introduction

In many empirical applications, e.g. estimation of mixed logit models, the researcher is faced with the problem that standard maximum likelihood estimation can not be applied, as analytical integration is not possible. Instead, methods such as quadrature or simulation are required for approximation of the integral. We suggest a Stata routine for multinomial logit models with unobserved heterogeneity using maximum simulated likelihood (MSL).[2] The purpose of this paper is twofold. First, we provide a description of the technical implementation of the estimation routine and discuss its properties. Further, we compare our estimation routine with the program `gllamm` which is implemented in Stata. `gllamm` is a very flexible programm incorporating a variety of multilevel models including mixed logit, see Rabe-Hesketh, Skrondal, and Pickles (2004) or Skrondal and Rabe-Hesketh (2005). Our routine differs from `gllamm` for computational reasons: whereas in `gllamm`, integrals are solved using classical Gauss Hermite or adaptive quadrature, we suggest simulation based on Halton sequences for integration.

[1]The following analysis is based on joint work with Peter Haan (Haan and Uhlendorff, 2006).

[2]Our approach closely follows Train (2003). Train implemented a program for mixed logit models in GAUSS.

In our analysis, we compare the performance of the estimation techniques using multilevel data about schooling from the `gllamm` manual.

Our empirical findings show that when the integral is reasonably well approximated the estimation techniques lead to nearly the same results. The advantage of Halton based simulation over classical Gauss Hermite quadrature is computational time; this advantage increases with the dimensions of the integral. Adaptive quadrature leads to more stable results relative to the other integration methods. However, again simulation is more time efficient. We find that maximum simulated likelihood leads to estimation results with reasonable accuracy in roughly half the time required when using adaptive quadrature.

The paper is organized as follows. In the next section, we provide a brief discussion about the estimation of multinomial logit models with unobserved heterogeneity using MSL. Hereafter, we present a description of the technical implementation of the estimation routine and discuss its properties. In section 4, we compare the performance of MSL with estimation based on classical and adaptive quadrature using multilevel data about schooling. The final section concludes.

A.2 Multinomial logit models with unobserved heterogeneity

Mixed logit models are a highly flexible class of models approximating any random utility model Train (2003). In this application we focus on a specific model of this broad class, the multinomial logit panel data model with random intercepts.[3] The results we present can be generalized and extended to other mixed logit models both with panel and cross-sectional data.

The theoretical framework of multinomial logit models can described as follows. Each individual i is faced with J different choices at time t. The individual receives a certain level of utility at each choice alternative and chooses the alternative that maximizes the utility. As well documented in the literature, see

[3]Note, we use panel data and multilevel data exchangeably.

e.g. Train (2003), the probability of making choice j conditional on observed characteristics X_{it} that vary between individuals and over time and unobserved individual effects α_i that are time constant has the following form:

$$P(j|X_{it}, \alpha_i) = \frac{exp(X_{it}\beta_j + \alpha_{ij})}{\Sigma_{k=1}^{J} exp(X_{it}\beta_k + \alpha_{ik})}$$

As the choice probabilities are conditioned on α_i, one must integrate over the distribution of the unobserved heterogeneity. Thus, the sample likelihood for the multinomial logit with random intercepts has the following form:

$$L = \prod_{i=1}^{N} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \prod_{j=1}^{J} \left(\frac{exp(X_{it}\beta_j + \alpha_j)}{\Sigma_{k=1}^{J} exp(X_{it}\beta_k + \alpha_k)} \right)^{d_{ijt}} f(\alpha) d\alpha \qquad \text{(A.1)}$$

where d_{ijt}=1 if individual i chooses alternative j at time t and zero otherwise. The coefficient vector and the unobserved heterogeneity term of one category are set to 0 for identification of the model. For convenience, we assume throughout our analysis that the unobserved heterogeneity α is identically and independently distributed over the individuals and follows a multivariate normal distribution with mean a and variance-covariance matrix **W**, $\alpha \sim f(a, W)$. In most applications, α is specified to be normally distributed. However, as Train (2003) points out, the distributional assumption depends on the research question; if more appropriate, distributions such as log-normal or uniform can be assumed. As standard in random-effects models, the unobserved heterogeneity α is required to be independent of the explanatory variables X_{it}.

To maximize the sample likelihood, one must integrate over the distribution of unobserved heterogeneity. Yet, there exists no analytical solution for the integral in equation (A.1). In the literature, many methods for integral approximation have been suggested and discussed. We focus on classical Gauss Hermite quadrature, adaptive quadrature, and simulation based on Halton sequences.

Gauss Hermite and adaptive quadrature

Gauss Hermite and adaptive quadrature are discussed in detail in the work of Rabe-Hesketh, Skrondal, and Pickles (2002). Gauss Hermite quadrature approximates an integral by a specified number of discrete points. Adaptive quadrature uses Bayes' rule to find quadrature weights that lead to a better approximation of the integral than those of normal Gauss Hermite quadrature, significantly increasing the accuracy of integration. The Stata program `gllamm` incorporates both integration methods, yet adaptive quadrature is strongly recommended for its higher accuracy (Rabe-Hesketh, Skrondal, and Pickles, 2002).

Estimation with Maximum Simulated Likelihood

We suggest integrating over the unobserved heterogeneity by using simulation and maximizing a simulated likelihood . MSL draws R values from the distribution of the unobserved heterogeneity with variance-covariance matrix $\mathbf{W}$. For each of these draws, the likelihood is calculated and then averaged over the R draws, which implies that instead of the exact likelihood a simulated sample likelihood (SL) is maximized:[4]

$$SL = \prod_{n=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} \prod_{j=1}^{J} \left(\frac{exp(X_{it}\beta_j + \alpha_j^r)}{\Sigma_{k=1}^{j} exp(X_{it}\beta_k + \alpha_k^r)} \right)^{d_{ijt}} \tag{A.2}$$

Consider an example with three different choices ($j = 3$). For identification, β_1 and α_{i1} are normalized to zero. We assume that the unobserved heterogeneity differs between the two other choices ($\alpha_{i2} \neq \alpha_{i3}$) and allow for correlation of these terms. Hence, the distribution of the unobserved heterogeneity can be described by a bivariate normal distribution with the following:

[4] When using random draws MSL is equivalent to the ML estimator if $N^{0.5}/R \to 0$ and both N and R$\to \infty$. For more detailed information, see e.g. Cameron and Trivedi (2005) .

$$\alpha \sim f\left(\begin{pmatrix} a_2 \\ \\ a_3 \end{pmatrix}, \begin{pmatrix} var_2 & cov_{23} \\ \\ cov_{23} & var_3 \end{pmatrix}\right) \tag{A.3}$$

This equation implies that when applying MSL, an approximate two-dimensional integral is needed. Each draw r consists of two values, $(\epsilon_2, \epsilon_3)'$, which follow a standard normal distribution. We apply a Cholesky decomposition of the variance-covariance matrix **W**. A Cholesky factor **L** of matrix **W** is defined such that **LL' =W**. Then the unobserved effects α^r are calculated by $\alpha^r = \mathbf{L}\epsilon^r$, which for our example implies the following:

$$\begin{pmatrix} \alpha_2 \\ \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 \\ \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} \epsilon_2 \\ \\ \epsilon_3 \end{pmatrix} \tag{A.4}$$

The example can be easily extended to more complex choice situations. However, with increasing number of choices integration becomes more and more time intensive as the dimension of the integral increases.

Instead of using random draws to obtain $(\epsilon_2, \epsilon_3)'$ we follow Train (2003) and recommend basing simulation on Halton sequences. Halton sequences generate quasirandom draws that provide a more systematic coverage of the domain of integration than independent random draws and induce a negative correlation over observations. Several studies such as Train (2000) and Bhat (2001) have shown that for mixed logit models, the accuracy can be markedly increased by using Halton sequences; the authors find in their studies that the results are more precise with 100 Halton draws than with $1,000$ random draws. These results confirm that quasirandom sequences go along with a lower integration error and faster convergence rates and therefore clearly require fewer draws than pseudorandom sequences.[5] However, as Train (2003) points out the use of Halton draws in simu-

[5]The expected integration error using pseudo-random sequences is of order $R^{-.5}$, whereas the theoretical upper bound for the integration error using quasi-random sequences is of order R^{-1}; see Bhat (2001) or Cameron and Trivedi (2005). This implies that a 10-fold increase in the number of quasirandom draws leads to the same improvement of accuracy as a 100-fold increase in the number of pseudorandom draws.

lation based estimation is not completely understood and caution is required. He provides an example of Halton sequences and discusses advantages and anomalies of this method in the context of mixed logit models. Computational time and estimation results slightly vary with the chosen primes for the Halton draws. This fact is documented by Train (2003), who found that the choice of the primes might noticeably affect the estimated coefficients.

The advantages of Halton draws might not hold for other models in the same way, see for example Cappellari and Jenkins (2006), who discuss Halton sequences for multivariate probit models.

A.3 Stata Routine for MSL estimation

In this section we provide an `ml model` statement that refers to a multinomial logit panel data model with two potentially correlated random intercepts that follow a bivariate normal distribution. This example can easily be extended to models with more alternatives.

For illustration, we apply our program to a real data set about teachers' evaluation of pupils behaviour.[6] The variables id and scy3 identify pupils and schools, respectively. Teachers group pupils in three different quality levels (tby) which is the dependent variable in our estimation. The data contain several additional variables explaining the quality level of the pupils like sex and provide information about 1313 pupils in 48 schools. The number of pupils differ between schools, i.e. we have an unbalanced panel.

The panel dimension of the data is not over time but over the pupils of a certain school (scy3). Hence, in the estimation, we can control for unobserved school-specific effects, but we do not control for individual-specific unobserved heterogeneity.[7] For simplicity, we condition the rating of teachers next to unobservable effects on only one observable variable, namely, sex.

[6] The data set is available as an ASCII file jspmix.dat (http://www.gllamm.org/jspmix.dat).

[7] The presented routine can easily be transferred to a model with time constant individual specific effects. Here the school (scy3) corresponds to the individual and one pupil to one individual observation at time t.

Before executing our program for MSL estimation, we apply the program mdraws by Cappellari and Jenkins (2006) to generate Halton Sequences and calculate the corresponding values following a standard normal distribution. mdraws can be used to create pseudonormal draws.

For each draw, the values (random_1‘r’ and random_2‘r’) must be th same for one observation within each unit, here within each school. Therefore, we create draws for every school and merge these draws to every pupil within each school. Here we approximate the integral by using 50 draws from the Halton sequence. We specify the primes used to create the Halton sequences as 7 and 11, because we later fit models with 150 draws and the number of draws should not be an integer multiple of any of the used primes. See Cappellari and Jenkins (2006) for details. We make use of the burn() option to drop the first 15 draws of each sequence because the initial elements of any two sequences can be highly correlated.

```
. matrix p = (7, 11)
. global draws "50"
. infile scy3 id sex stag ravi fry3 tby using jspmix.dat, clear
(1313 observations read)
. save jspmix.dta, replace
. keep scy3
. sort scy3
. by scy3: keep if _n==1
(1265 observations deleted)
. mdraws, neq(2) dr($draws) prefix(c) burn(15) prime (p)
Created 50 Halton draws per equation for 2 equations. Number of
initial draws dropped per equation = 15 . Primes used:
   7  11
. forvalues r=1/$draws{
  2. gen random_1‘r’=invnorm(c1‘r’)
  3. gen random_2‘r’=invnorm(c2‘r’)
  4.          }
. sort scy3
. save mdraws_${draws}, replace
file mdraws_50.dta saved
. use "jspmix.dta",clear
. sort scy3
. merge scy3 using mdraws_${draws}.dta
variable scy3 does not uniquely identify observations in the master
data
. drop _merge
. sort scy3
```

To get appropriate starting values for the coefficient vector, we use `mlogit` to estimate a multinomial logit model without random intercepts. The variables a1, a2, and a3 take on the value 1 if the choice 1, 2 or 3 is made, respectively, and zero otherwise; the variable are defined using the `tabulate` command.

```
. mlogit tby sex, base(1)
 (output omitted)
. matrix Init= e(b)
. tab tby, gen(a)
 (output omitted)
. sort scy3
```

The following `ml model` statement can be applied independently of the chosen type of draws (e.g., pseudorandom or Halton). We apply the method d0 because we fit panel data models with joint unobserved heterogeneity for groups of observations. The method d0 requires the researcher to supply the log-likelihood function. The first and second derivatives are obtained numerically; i.e., one need not supply analytically calculations of the gradient and the Hessian of the log-likelihood function.[8]

```
program define mlogit_sim_d0
  args todo b lnf
  tempvar etha2 etha3 random1 random2 lj pi1 pi2 pi3 sum lnpi L1 L2 last
  tempname lnsig1 lnsig2 atrho12 sigma1 sigma2 cov12

  mleval ‘etha2’   = ‘b’, eq(1)
  mleval ‘etha3’   = ‘b’, eq(2)
  mleval ‘lnsig1’  = ‘b’, eq(3) scalar
  mleval ‘lnsig2’  = ‘b’, eq(4) scalar
  mleval ‘atrho12’ = ‘b’, eq(5) scalar

  qui {
    scalar ‘sigma1’=(exp(‘lnsig1’))^2
    scalar ‘sigma2’=(exp(‘lnsig2’))^2
    scalar ‘cov12’=[exp(2*‘atrho12’)-1]/[exp(2*‘atrho12’)+1]*///
       (exp(‘lnsig2’))*(exp(‘lnsig1’))
    gen double ‘random1’ = 0
    gen double ‘random2’ = 0
    gen double ‘lnpi’=0
    gen double ‘sum’=0
    gen double ‘L1’=0
```

[8] The principles of computing maximum likelihood estimators with Stata are described in Gould, Pitbaldo, and Sribney (2003).

```
    gen double ‘L2’=0
    by scy3: gen byte ‘last’=(_n==_N)
    gen double ‘pi1’= 0
    gen double ‘pi2’= 0
    gen double ‘pi3’= 0
  }
  matrix W = ( ‘sigma1’ , ‘cov12’ \ ‘cov12’ , ‘sigma2’)

  capture matrix L=cholesky(W)

  if _rc != 0 {
        di "Warning: cannot do Cholesky factorization of rho matrix"
    }

  local l11=L[1,1]
  local l21=L[2,1]
  local l22=L[2,2]

  local repl=${draws}
  local r=1
  forvalues r=1/$draws{
   qui {
    replace ‘random1’ = random_1‘r’*‘l11’
    replace ‘random2’ = random_2‘r’*‘l22’ + random_1‘r’*‘l21’

    replace ‘pi1’= 1/(1 + exp(‘etha2’+‘random1’)+exp(‘etha3’+‘random2’))
    replace ‘pi2’= exp(‘etha2’+‘random1’)*‘pi1’
    replace ‘pi3’= exp(‘etha3’+‘random2’)*‘pi1’

    replace ‘lnpi’=ln(‘pi1’*a1+‘pi2’*a2+‘pi3’*a3)

    by scy3: replace ‘sum’=sum(‘lnpi’)
    by scy3: replace ‘L1’ =exp(‘sum’[_N]) if _n==_N

    by scy3: replace ‘L2’=‘L2’+‘L1’ if _n==_N
   }
  }

  qui gen ‘lj’=cond(!‘last’,0, ln(‘L2’/‘repl’))
  qui mlsum ‘lnf’=‘lj’
  if (‘todo’==0|‘lnf’>=.) exit

end
```

Instead of estimating the variances and the correlation coefficient directly, we estimate transformed variables of these parameters, i.e. the logarithm of the standard deviations (lnsig1 and lnsig2) and the inverse hyperbolic tangent of

ρ (atrho12), to constrain them within their valid limits. Therefore, the first step in our program is to calculate the variances (sigma1 and sigma2) and the covariance (cov12) of the bivariate normal distribution. Then we apply a Cholesky decomposition of the covariance matrix $\mathbf{W}$. To do this, the matrix $\mathbf{W}$ has to be positive definite at each iteration. If not, our program traps the error, shows a warning, and uses the most recent estimate of $\mathbf{W}$ which is guaranteed to be positive definite. This is assured by the command `capture`.[9]

We calculate the likelihood for each draw based on the individual specific quasirandom terms random1 and random2 within the following loop. The two terms random1_‘r’ and random2_‘r’ are multiplied with the elements of the Cholesky matrix $\mathbf{L}$, following equation (A.4). The probabilities of making choice 1, 2 or 3 are expressed by pi1, pi2 and pi3. With the information about the realized choices, captured in variables a1, a2 and a3, the likelihood is evaluated for each observation. The corresponding log likelihood values are added up within each unit for each draw (sum) and this sum is exponentiated for the last observation per unit (L1). These likelihood values are added up over all draws (L2). Following equation (A.2), the approximated likelihood is the average over the r draws. The simulated likelihood can be maximized using the options to the `ml maximize` and `ml model` commands. To set the starting values, we use the command `ml init`. For the β, we use the estimated coefficients from the `mlogit` saved as matrix Init. The starting values of lnsig1, lnsig2 and atrho12 are set to 0.5.

```
. ml model d0 mlogit_sim_d0 ( tby = sex) ( tby = sex)  /lnsig1 /lnsig2 /atsig12
. matrix start = (Init)
. ml init start   0.5 0.5 0.5, copy
. ml maximize

initial:       log likelihood = -1338.0475
rescale:       log likelihood = -1338.0475
rescale eq:    log likelihood = -1301.4639
Iteration 0:   log likelihood = -1301.4639
Iteration 1:   log likelihood = -1300.4893
Iteration 2:   log likelihood = -1299.4587
Iteration 3:   log likelihood = -1299.4509
```

[9] The procedure is the same as in the program `mvprobit` by Cappellari and Jenkins (2003).

```
Iteration 4:   log likelihood = -1299.4509

                                                  Number of obs   =       1313
                                                  Wald chi2(1)    =      14.22
Log likelihood = -1299.4509                       Prob > chi2     =     0.0002

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
eq1          |
         sex |   .5488225     .14552     3.77   0.000     .2636085    .8340364
       _cons |     .59589   .1394991     4.27   0.000     .3224768    .8693032
-------------+----------------------------------------------------------------
eq2          |
         sex |   1.104577   .1748037     6.32   0.000     .7619681    1.447186
       _cons |  -.5663381   .1816152    -3.12   0.002    -.9222974   -.2103788
-------------+----------------------------------------------------------------
lnsig1       |
       _cons |  -.3369519   .1695314    -1.99   0.047    -.6692274   -.0046763
-------------+----------------------------------------------------------------
lnsig2       |
       _cons |  -.1021489   .1602249    -0.64   0.524    -.4161839    .2118861
-------------+----------------------------------------------------------------
atsig12      |
       _cons |   1.614593   .3185383     5.07   0.000     .9902697    2.238917
------------------------------------------------------------------------------

. _diparm lnsig1, function((exp(@))^2) ///
>     deriv(2*(exp(@))*(exp(@))) label("sigma1")
      sigma1 |   .5097149   .1728254                      .2622506    .9906909

. _diparm lnsig2, function((exp(@))^2) ///
>     deriv(2*(exp(@))*(exp(@))) label("sigma2")
      sigma2 |   .8152196   .2612369                       .435018    1.527713

. _diparm atsig12, tanh label("roh12")
       roh12 |   .9238359   .0466745                      .7574773    .9775391

. . _diparm atsig12 lnsig1 lnsig2, //
> function([exp(2*@1)-1]/[exp(2*@1)+1]*(exp(@2))*(exp(@3))) ///
> deriv(-(2*exp(2*@1+@2+@3)*(-1+exp(2*@1))/(1+exp(2*@1))^2)+//
>      2*exp(2*> @1+@2+@3)/(1+exp(2*@1)) ///
>      [exp(2*@1)-1]/[exp(2*@1)+1]*(exp(@2))*(exp(@3)) ///
>      [exp(2*@1)-1]/[exp(2*@1)+1]*(exp(@2))*(exp(@3))) label("cov12")
       cov12 |   .5955193    .188545                      .2259779    .9650606
```

As mentioned above, we estimate the variances and the covariance in a transformed metric. We make use of the program `_diparm` to calculate and display the parameters and their standard errors after the estimation. For this task, we must calculate the first derivative of the function. Also we can use `_diparm` to calculate the correlation and its standard error.

A.4 Illustrations

In the following, we discuss the empirical performance of the MSL routine using a multilevel data set about schooling (Junior School Project) that is taken from the `gllamm` manual (Rabe-Hesketh, Skrondal, and Pickles, 2004). We described the data in the previous section. The main purpose of this illustration is to provide a comparison of the above described integration methods, Gauss Hermite and adaptive quadrature using `gllamm` and simulation based on Halton draws using our MSL routine. We are interested in two findings: 1) the accuracy of the procedures, evaluated for the stability of estimation results and 2) the computational time they require. Further, we want to show how the two estimators perform when the dimension of the integrals increases. Therefore, we fit models with only one random term (one-dimensional integral) and with two random terms (two-dimensional integral). One random term implies that unobserved effects are constant between the alternatives. In the second example (two random terms), the heterogeneity varies between the alternatives and is potentially correlated. The structure of unobserved heterogeneity is the same as in the example described in section 2.

Computational time and accuracy of integral approximation depend on the chosen number of quadrature points or number of draws when estimating. Therefore, we present several estimations by increasing the number of quadrature points and draws. As there is a trade off between accuracy of integration and computational time the number of points or draws can become a crucial variable. Providing a rigid test indicating the optimal number of draws is difficult. In practice, researchers often vary the number of draws or points to see whether the coefficients and the log likelihood remain constant as an indication whether an adequate number of draws is chosen (Cameron and Trivedi, 2005). We present results of six estimations using MSL with 25, 50, 100, 150, 200, and 500 draws from the Halton sequences and six estimations with Gauss Hermite and adaptive quadrature, both with 4, 8 and 16 points.[10] Note, as we do not directly test for accuracy, the comparison needs to be interpreted carefully. All estimates were computed with Intercooled Stata version 8.2 on a 3GHz Pentium 4 PC running Windows 2000 Professional. To make computational time between both methods comparable, we used the same starting values for all estimations.

In the following, we present the `gllamm` command for estimation of the model with the two dimensional integral using four quadrature points (Gauss Hermite). For further description of the syntax, see Rabe-Hesketh, Skrondal, and Pickles (2004).

```
use "jspmix.dta",clear
mlogit tby sex, base(1)
matrix Init= e(b)
scalar var = exp(0.5)
matrix start= Init, var, var, 0.5
matrix colnames start= sex _cons sex _cons a2 a3 _cons
matrix coleq start= c2 c2  c3 c3  scy1_1  scy1_2 scy1_2_1
gen school =scy3 sort school sex tby
gen patt =_n
expand 3
sort patt
qui by patt: gen alt= _n
gen chosen =alt ==tby
sort pat alt
tab alt, gen (a)
gen dum=1
replace dum=0 if a1==1
eq dum: dum
eq a2: a2
eq a3: a3

gllamm alt sex, expand(patt chosen m)i(scy3)link(mlogit) /*
*/family(binom) nrf(2) eq(a2 a3) nip(4) trace from(start)
```

[10] In addition to that we estimated the model using MSL based on pseudorandom draws. Our results are in line with previous studies, e.g. Train (2000) and Bhat (2001), and indicate that many more pseudorandom draws than Halton draws are required to get relatively stable results.

Table A.1: One random intercept: Maximum Simulated Likelihood

Parameter	Coefficient (SE)					
tby= 2						
sex	0.543	0.551	0.549	0.550	0.550	0.550
	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)
constant	0.685	0.598	0.592	0.592	0.592	0.591
	(0.141)	(0.145)	(0.146)	(0.145)	(0.146)	(0.145)
tby= 3						
sex	1.064	1.072	1.070	1.071	1.071	1.070
	(0.171)	(0.171)	(0.171)	(0.171)	(0.171)	(0.171)
constant	-0.399	-0.486	-0.492	-0.492	-0.492	-0.493
	(0.160)	(0.163)	(0.164)	(0.164)	(0.164)	(0.164)
lnsig1	-0.391	-0.289	-0.301	-0.301	-0.321	-0.312
	(0.146)	(0.154)	(0.155)	(0.159)	(0.163)	(0.162)
sig1	0.457	0.561	0.547	0.548	0.526	0.536
	(0.133)	(0.172)	(0.170)	(0.174)	(0.172)	(0.173)
Log likelihood	-1303.791	-1303.605	-1303.751	-1303.937	-1303.658	-1303.740
Time (hh:mm:ss)	00:00:21	00:00:41	00:01:25	00:02:10	00:03:10	00:08:31
No. of Draws	25	50	100	150	200	500

Numbers of Observations: 1,313.
Source: http://www.gllamm.org/jspmix.dat

Table A.1 shows the MSL results for the model with a common term of unobserved heterogeneity. Comparing the coefficients and the log likelihood between the estimations, we find that the results are fairly stable when using at least 50 draws. When using only 25 Halton draws, the deviations of the coefficients from those obtained with better approximated integrals can be seen. However, even with more than 100 draws, we find that results slightly differ between the number of draws; the log likelihood varies between the estimations in the first decimal place. Estimation time varies between the estimations with an acceptable approximation of the integral from 42 seconds (50 draws) to 8 minutes and 31 seconds (500 draws); estimation results suggest that computational time increases approximately linear with the number draws.

Table A.2 compares one random intercept calculated with both Gauss Hermite and adaptive quadrature. Comparing the results derived with simulation with those estimated with quadrature, we find that the estimation results are quite similar when the integral is reasonably well approximated. When using

Table A.2: One random intercept: Gauss Hermite and Adaptive Quadrature

Parameter	Coefficient (SE)					
tby= 2						
sex	0.553	0.554	0.549	0.550	0.550	0.550
	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)
constant	0.693	0.619	0.593	0.594	0.594	0.594
	(0.146)	(0.155)	(0.147)	(0.145)	(0.146)	(0.146)
tby= 3						
sex	1.074	1.075	1.070	1.071	1.071	1.070
	(0.171)	(0.171)	(0.171)	(0.171)	(0.171)	(0.171)
constant	-0.391	-0.465	-0.492	-0.490	-0.491	-0.491
	(0.165)	(0.172)	(0.166)	(0.163)	(0.164)	(0.164)
sig1	0.398	0.564	0.530	0.551	0.543	0.544
	(0.101)	(0.181)	(0.166)	(0.178)	(0.175)	(0.175)
Log likelihood	-1305.189	-1303.681	-1303.843	-1303.802	-1303.804	-1303.804
Time (hh:mm:ss)	00:00:21	00:00:46	00:01:10	00:01:24	00:01:42	00:03:12
No. of quadrature points	4	8	16	4 (Adaptive)	8 (Adaptive)	16 (Adaptive)

Numbers of Observations: 1313.
Source: http://www.gllamm.org/jspmix.dat

Gauss Hermite quadrature, at least 8 quadrature points are required for integration. The log likelihood and the coefficients clearly differ between the estimation with 4 and 8 points.

Turning to the adaptive quadrature, the picture changes. With only four quadrature points, the integral seems to be reasonably well approximated, as a further increase in quadrature points leads to very similar estimated parameters. This finding underscores the result of Rabe-Hesketh, Skrondal, and Pickles (2002) who show the computational advantage of the adaptive quadrature versus Gauss Hermite quadrature.

For the one-dimensional integral, Halton based simulation performs similarly to quadrature. Relative to Gauss Hermite quadrature, we find hardly any difference in computational time for a comparable degree of accuracy. The adaptive quadrature leads to more stable results with 4 quadrature points, computation time, however, is higher than in a simulation with 50 draws and about the same as in a simulation with 100 draws. This finding indicates that with one term there is no advantage of using MSL relative to adaptive quadrature.

Table A.3: Two random intercepts: Maximum Simulated Likelihood

Parameter	Coefficient (SE)					
tby= 2						
sex	0.542	0.549	0.546	0.545	0.546	0.546
	(0.145)	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)
constant	0.616	0.596	0.577	0.601	0.576	0.593
	(0.142)	(0.139)	(0.144)	(0.140)	(0.142)	(0.141)
tby= 3						
sex	1.095	1.105	1.099	1.102	1.101	1.101
	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)
constant	-0.534	-0.566	-0.585	-0.563	-0.585	-0.569
	(0.184)	(0.182)	(0.178)	(0.180)	(0.181)	(0.180)
lnsig1	-0.367	-0.337	-0.327	-0.366	-0.362	-0.361
	(0.201)	(0.170)	(0.174)	(0.167)	(0.175)	(0.171)
lnsig2	-0.153	-0.102	-0.145	-0.142	-0.162	-0.158
	(0.167)	(0.160)	(0.158)	(0.154)	(0.163)	(0.161)
athro	1.535	1.615	1.471	1.550	1.487	1.496
	(0.422)	(0.319)	(0.320)	(0.339)	(0.353)	(0.346)
sig1	0.479	0.510	0.520	0.481	0.484	0.485
	(0.192)	(0.173)	(0.181)	(0.160)	(0.170)	(0.166)
sig2	0.735	0.815	0.749	0.753	0.724	0.729
	(0.246)	(0.261)	(0.236)	(0.231)	(0.236)	(0.234)
cov12	0.54	0.596	0.561	0.550	0.535	0.538
	(0.185)	(0.189)	(0.184)	(0.172)	(0.181)	(0.177)
cor	0.911	0.924	0.900	0.914	0.903	0.904
	(0.071)	(0.047)	(0.061)	(0.056)	(0.065)	(0.063)
Log likelihood	-1,299.900	-1,299.451	-1,299.700	-1,299.635	-1,299.726	-1,299.599
Time (hh:mm:ss)	00:00:45	00:01:34	00:03:34	00:05:00	00:06:52	00:19:54
No. of Draws	25	50	100	150	200	500

Numbers of Observations: 1313.
Source: http://www.gllamm.org/jspmix.dat

In the following discussion, the complexity of the estimation increases by allowing the unobserved heterogeneity to differ between the alternatives. Here the advantage of computational time of Halton based simulation over Gauss Hermite quadrature becomes evident. As shown in Table A.3, with at least 100 draws, coefficients and the log likelihood become relatively stable. For 100 draws, the estimation takes more than 3 minutes. Table A.4 shows that for a comparable level of integral approximation, Gauss Hermite quadrature requires more than

Table A.4: Two random intercepts: Gauss Hermite and Adaptive Quadrature

Parameter	Coefficient (SE)					
tby= 2						
sex	0.548	0.551	0.546	0.547	0.546	0.546
	(0.145)	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)
constant	0.668	0.621	0.595	0.598	0.597	0.597
	(0.142)	(0.142)	(0.141)	(0.140)	(0.141)	(0.141)
tby= 3						
sex	1.104	1.105	1.101	1.102	1.101	1.101
	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)
constant	-0.480	-0.539	-0.567	-0.564	-0.565	-0.565
	(0.181)	(0.181)	(0.181)	(0.180)	(0.180)	(0.180)
sig1	0.352	0.504	0.480	0.489	0.488	0.488
	(0.098)	(0.169)	(0.168)	(0.171)	(0.170)	(0.170)
sig2	0.596	0.752	0.730	0.743	0.739	0.738
	(0.169)	(0.238)	(0.234)	(0.240)	(0.238)	(0.238)
cov12	0.406	0.560	0.537	0.547	0.545	0.545
	(0.108)	(0.180)	(0.177)	(0.182)	(0.181)	(0.181)
cor	0.887	0.910	0.907	0.908	0.908	0.908
	-	-	-	-	-	-
Log likelihood	-1300.950	-1299.482	-1299.681	-1299.663	-1299.664	-1299.665
Time (hh:mm:ss)	00:02:47	00:11:38	00:47:41	00:08:16	00:30:38	02:03:12
No. of quadrature points	4	8	16	4 (Adaptive)	8 (Adaptive)	16 (Adaptive)

Numbers of Observations: 1313.

Source: http://www.gllamm.org/jspmix.dat

11.5 minutes. Results from MSL become more stable with 200 and 500 draws. The estimation with 200 draws takes less than 7 minutes, and the one with 500 draws about 20 minutes. When doubling the number of quadrature points for the Gauss hermite approach, computational time approximately quadruples (50 minutes) and the results are similar to those from the adaptive quadrature.

With adaptive quadrature, again 4 points are sufficient for approximation of the integral. Results hardly change with more quadrature points. Computational time with four points is about 8 minutes. Relative to simulation, adaptive quadrature leads to more robust results. However, using simulation with 100 draws, one can approximate the integral such that coefficients and the log likelihood are approximately stable in less than 3.5 minutes. Here the trade off between computational time and accuracy becomes evident. Halton based simulation leads to results in less computational time, whereas adaptive quadrature provides results that are more stable.

From a practical point of view, the implementation of MSL based on Halton sequences is relatively simple and has significant advantages in computational time if it is compared to Gauss Hermite quadrate and simulation based on pseudorandom sequences, not reported here. This implementation is particularly true for higher-dimensional integrals. Compared with adaptive quadrature, our routine seems to be less stable. However, given the advantage of computational time, Halton based MSL could be the adequate model choice. The time advantage becomes even more important when sample size or the dimension of the integral increases.[11]

Therefore, we recommend the presented routine as an alternative to the quadrature approach implemented in `gllamm`. Moreover, the principles of our routine can be a useful starting point for evaluating likelihood functions which are not pre-programmed in Stata and involve a multivariate normal distribution of the unobserved heterogeneity.

A.5 Conclusion

In this paper we have suggested a Stata routine for multinomial logit models with unobserved heterogeneity using MSL based on Halton sequences. The routine refers to a model with two random intercepts but can easily be extended to models with a higher dimension. Further extensions of the presented code are possible. One example is Haan (2005), fitting a dynamic conditional logit model. Another example is the dynamic multinomial logit model with endogenous panel attrition, estimated in Chapter 4.

Using multilevel data about schooling we compare the performance of our code to that of the Stata program `gllamm`, which numerically approximates integrals using classical Gauss Hermite quadrature and adaptive quadrature. Esti-

[11] Using Bayes' rule for simulation might be one way to reduce the tradeoff between estimation time and accuracy. Train (2003) suggests to employ Bayesian simulation instead of classical MSL, as the Bayesian method leads to consistent estimates even with a fixed number of draws.

mation by MSL provides approximately the same estimation results as estimation with Gauss Hermite quadrature or adaptive quadrature. Compared with classical quadrature, simulation markedly reduces computational time when a higher dimensional integral needs to be approximated. However, relative to the adaptive quadrature, the advantage of simulation vanishes in our example. Adaptive quadrature leads to very stable results with only a few quadrature points (four points). Estimations with 100 draws are less stable but lead to qualitatively the same results and take roughly half the estimation time. This finding underscores the tradeoff between computational time and accuracy of the results, which becomes important if estimation takes not a few but instead hours or days.

Bibliography

ARULAMPALAM, W., A. L. BOOTH, AND M. P. TAYLOR (2000): "Unemployment Persistence," *Oxford Economic Papers*, 52, 24–50.

ASPLUND, R., P. J. SLOANE, AND I. THEODOSSIU (1998): *Low Pay and Earnings Mobility in Europe*. Edward Elgar, Chelthenham.

BARRETT, G. F. (2000): "The effect of educational attainment on welfare dependence: Evidence from Canada," *Journal of Public Economics*, 77, 209–232.

BAUER, T., B. DIETZ, K. F. ZIMMERMANN, AND E. ZWINTZ (2005): "German Migration: Development, Assimilation, and Labour Market Effects," in *European Migration: What Do We Know?*, ed. by K. F. Zimmermann, pp. 197–261. Oxford University Press, Oxford, New York.

BELZIL, C. (2001): "Unemployment insurance and subsequent job duration: job matching vs. unobserved heterogeneity," *Journal of Applied Econometrics*, 16, 619–633.

BENDER, S., AND W. KARR (1993): "Arbeitslosigkeit von ausländischen Arbeitnehmern," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 26, 192–206.

BHAT, C. (2001): "Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model," *Transportation Research B*, 35, 677–693.

BLANK, R. M. (1989): "Analyzing the Length of Welfare Spells," *Journal of Public Economics*, 39, 245–273.

Blundell, R., and T. MaCurdy (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics, Vol. 3A*, ed. by O. Ashenfelter, and D. Card, pp. 1559–1695. Elsevier, Amsterdam.

Bonin, H., and H. Schneider (2006): "Workfare: Eine wirksame Alternative zum Kombilohn," *Wirtschaftsdienst*, 86, 645–650.

Borjas, G. J. (1994): "The economics of immigration," *Journal of Economic Literature*, 32, 1667–1717.

Boss, A. (2002): *Sozialhilfe, Lohnabstand und Leistungsanreize. Empirische Analyse für Haushaltstypen und Branchen in West- und Ostdeutschland.* Kieler Studien 318, Berlin.

Breuer, W., and D. Engels (2003): *Grundinformationen und Daten zur Sozialhilfe. Im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung.*

Bundesagentur für Arbeit (2006): *Analytikreport der Statistik. Analyse des Arbeitsmarktes für Ausländer. Dezember 2006.*

Cameron, C., and P. K. Trivedi (2005): *Microeconometrics. Methods and Applications.* Cambridge University Press, New York.

Cappellari, L. (2002): "Do the 'working poor' stay poor? An analysis of low pay transitions in Italy," *Oxford Bulletin of Economics and Statistics*, 64, 87 – 110.

Cappellari, L., and S. P. Jenkins (2003): "Multivariate probit regression using simulated maximum likelihood," *The Stata Journal*, 3, 278–294.

——— (2004): "Modelling Low Pay Transition Probabilities, Accounting for Panel Attrition, Non-Response, and Initial Conditions," *CESifo Working Paper*, 1232.

——— (2006): "Calculation of multivaraite normal probabilities by simulation, with applications to maximum simulated likelihood estimation," *Stata Journal*, 6, 156–189.

Cramer, U. (1984): "Multivariate Analyse von Arbeitslosigkeitsquoten," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 17, 330–335.

Devine, T. J., and N. M. Kiefer (1991): *Empirical Labor Economics. The Search Approach.* New York, Oxford.

Dustmann, C. (1993): "Earnings adjustment of temporary migrants," *Journal of Population Economics*, 6, 153–168.

Engels, D. (2001): "Abstand zwischen Sozialhilfe und unteren Arbeitnehmereinkommen: Neue Ergebnisse zu einer alten Kontroverse," *Sozialer Forschritt*, 50, 56–62.

European Commission (2004): *Employment in Europe 2004. Recent Trends and Prospects. Chapter 4: Labour market transitions and advancement: temporary employment and low-pay in Europe.*

Fertig, M., and S. Schurer (2006): "Labour market performance of immigrants in Germay revisited: The importance of omitted variable and attrition bias," *mimio.*

Flaig, G., G. Licht, and V. Steiner (1993): "Testing for State Dependence Effects in a Dynamic Model of Male Unemployment Behavior," in *Panel Data And Labour Market Dynamics*, ed. by P. J. Henning Bunzel, and N. Westergard-Nielsen, pp. 114–178. North-Holland, Amsterdam.

Fortin, B., G. Lacroix, and S. Drolet (2004): "Welfare benefits and the duration of welfare spells: evidence from a natural experiment in Canada," *Journal of Public Economics*, 88, 1495–1520.

Frijters, P., M. A. Shields, and S. W. Price (2005): "Job Search Methods And Their Success: A Comparison of Immigrants and Natives in the UK," *The Economic Journal*, 115, F359–F376.

Gangl, M. (1998): "Sozialhilfebezug und Arbeitsmarktverhalten. Eine Längsschnittanalyse der Übergänge aus der Sozialhilfe in den Arbeitsmarkt," *Zeitschrift für Soziologie*, 27, 212–232.

Gebauer, R., H. Petschauer, and G. Vobruba (2002): *Wer sitzt in der Armutsfalls? Selbstbehauptung zwischen Sozialhilfe und Arbeitsmarkt.* Edition Sigma, Berlin.

Gittleman, M. (2001): "Declining Caseloads: What Do the Dynamics of Welfare Participation Reveal?," *Industrial Relations*, 40, 537–570.

Gong, X., A. van Soest, and E. Villagomez (2004): "Mobility in the Urban Labor Market: A Panel Data Analysis for Mexico," *Economic Development and Cultural Change*, 53 (1), 1–36.

Gould, W., J. Pitbaldo, and W. Sribney (2003): *Maximum Likelihood Estimation with Stata.* Stata Corporation, Texas.

Gourieroux, C., and A. Monfort (1993): "Simulation-based inference: A survey with special reference to panel data models," *Journal of Econometrics*, 59, 5–33.

Greene, W. H. (2003): *Econometric Analysis.* Prentice Hall, 5th edn.

Gregory, M., and R. Jukes (2001): "Unemployment and Subsequent Earnings: Estimating Scarring among British Men, 1984-94," *Economic Journal*, 111, FF607–25.

Görzig, B., M. Gornig, and A. Werwatz (2005): "Explaining Eastern Germany's Wage Gap: The Impact of Structural Change," *Post-Communist Economies*, 17, 449–464.

Haan, P. (2005): "State Dependence and Female Labor Supply in Germany: The Extensive and the Intensive Margin," *DIW Discussion Paper*, 538.

Haan, P., and A. Uhlendorff (2006): "Estimation of multinomial logit models with unobserved heterogeneity using maximum simulated likelihood," *Stata Journal*, 6, 229–245.

Haisken-DeNew, J., and J. Frick (2005): *Desktop Compendium to The German Socio-Economic Panel Study (SOEP)*. DIW, Berlin.

Hajivassiliou, V., and P. Ruud (1994): "Classical estimation methods for LDV models using simulation," in *Handbook of Econometrics*, ed. by R.Engle, and D. McFadden, pp. 2393–2441. North-Holland, Amsterdam.

Ham, J. C., and R. J. LaLonde (1996): "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training," *Econometrica*, 64, 175–205.

Han, A., and J. A. Hausman (1990): "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5, 1–28.

Haustein, T., and S. Krieger (2004): "Ergebnisse der Sozialhilfe- und Asylbewerberleistungsstatistik 2002," *Wirtschaft und Statistik*, pp. 192–208.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

Heckman, J. J. (1981a): "Heterogeneity and state dependence," in *Studies in Labor Markets*, ed. by S. Rosen, pp. 91–139. Chicago University Press, Chicago.

——— (1981b): "The Incedental Paramter Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. Manski, and D.McFadden, pp. 179–195. MIT Press, Cambridge, MA.

HECKMAN, J. J., AND G. J. BORJAS (1980): "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence," *Economica*, 47, 247–283.

HONORÉ, B. E., AND E. KYRIAZIDOU (2000): "Panel Data Discrete Choice Models with Lagged Dependent Variables," *Econometrica*, 68, 839–874.

HOYNES, H., AND T. MACURDY (1994): "Has the Decline in Benefits Shortened Welfare Spells?," *The American Economic Review*, 84, 43–48.

HYSLOP, D. R. (1999): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women," *Econometrica*, 67, 1255–1294.

JACOBSEN, L. S., R. J. LALONDE, AND D. G. SULLIVAN (1993): "Earnings Losses of Displaced Workers," *American Economic Review*, 83, 685–709.

JENKINS, S. P. (2004): "Independent competing risks models for interval-censored survival time data," *Unpublished Paper, ISER, University of Essex.*

JONES, M. K., R. J. JONES, P. D. MURPHY, AND P. J. SLOANE (2005): "The Dynamics of the National Minimum Wage: Transitions Between Different Labour Market States," *IZA Discussion-Paper*, 1690.

KEANE, M. P., AND K. I. WOLPIN (2002): "Estimating Welfare Effects Consistent With Forward-Looking Behavior: Part II: Empirical Results," *Journal of Human Resources*, 37, 600–622.

KETZLER, L. G. (1998): "Job Displacement," *Journal of Economic Perspectives*, 12, 115–136.

KOGAN, I. (2004): "Last Hired, First Fired? The Unemployment Dynamics of Male Immigrants in Germany," *European Sociological Review*, 20, 445–461.

LANG, G. (2005): “The Difference Between Wages and Wage Potentials: Earnings Disadvantages of Immigrants in Germany,” *Journal of Economic Inequality*, 3, 21–42.

LEMIEUX, T., AND K. MILLIGAN (2006): “Incentive Effects of Social Assistance: A Regression Discontinuity Approach,” *Analytical Studies - Research Paper Series*, 280.

MCCORMICK, B. (1990): “A Theory of Signalling During Job Search, Employment Efficiency, and “Stigamtized” Jobs,” *Review of Economic Studies*, 57, 299–313.

MEYER, B. D. (1990): “Unemployment Insurance and Unemployment Spells,” *Econometrica*, 58, 757–782.

MÜHLEISEN, M., AND K. F. ZIMMERMANN (1994): “A Panel Analysis of Job Changes and Unemployment,” *European Economic Review*, 38, 793–801.

MICHAUD, P.-C., AND K. TATSIRAMOS (2005): “Employment Dynamics of Married Women in Europe,” *IZA Discussion-Paper*, 1706.

MOFFITT, R. (1992): “Incentive Effects of the U.S. Welfare System: A Review,” *Journal of Economic Literature*, 30, 1–61.

——— (2002): “Welfare Programs and Labor Supply,” in *Handbook of Public Economics, Vol.4*, ed. by A. Auerbach, and M. Feldstein, pp. 2393–2430. Elsevier Science, Amsterdam.

NARENDRANATHAN, W., AND M. B. STEWART (1993a): “How Does the Benefit Effect Vary as Unemployment Spells Lengthen?,” *Journal of Applied Econometrics*, 8, 361–381.

——— (1993b): “Modeling the Probability of Leaving Unemployment: Competing Risks Models with Flexible Base-line Hazards,” *Applied Statistics*, 42, 63–83.

Ochel, W. (2003): "Welfare to Work in the U.S.: A Model for Germany?," *Finanzarchiv*, 59, 91–119.

OECD (2006): *Employment Outlook. Boosting Jobs And Incomes.* OECD, Paris.

Prowse, V. (2005): "State Dependence in a Multi-State Model of Employment Dynamics," *IZA Discussion-Paper*, 1623.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles (2002): "Reliable estimation of generalised linear mixed models using adaptive quadrature," *The Stata Journal*, 2, 1–21.

——— (2004): "GLLAMM Manual, Technical Report," *U.C. Berkeley Division of Biostatistics Working Paper Series*, Working Paper 160.

Reinberg, A., and M. Hummel (2005): "Höhere Bildung schützt auch in der Krise vor Arbeitslosigkeit," *IAB Kurzbericht*, 9/2005, 1–6.

Rhein, T., H. Gartner, and G. Krug (2005): "Niedriglohnsektor. Aufstiegschancen fuer Geringverdiener verschlechtert," *IAB Kurzbericht*, 3/2005, 1–5.

Riphahn, R. T. (1999): *Why Did Social Assistance Dependence Increase? - The Dynamics of Social Assistance Dependence and Unemployment in Germany.* unpublished Habilitationsschrift, University of Munich.

Roed, K., and M. Nordberg (2003): "Temporary layoffs and the duration of unemployment," *Labour Economics*, 10, 381–398.

Roed, K., and Zhang (2005): "Unemployment duration and economic incentives - a quasi random-assignment approach," *European Economic Review*, 49, 1799–1825.

Schmidt, C. M. (1997): "Immigrant Performance in Germany. Labor Earnings of Ethnic German Migrants and Foreign Guest-Workers.," *Quarterly Review of Economics and Finance*, 37, 379–397.

Schneider, H., and A. Uhlendorff (2004): “The Transition from Welfare to Employment and the Role of potential Labor Income,” *IZA Discussion-Paper*, 1420.

——— (2005): “Transition from Welfare to Employment: Does the Ratio between Labor Income and Social Assistance Matter?,” *Schmollers Jahrbuch*, 125, 51–61.

——— (2006): “Die Wirkung der Hartz-Reform im Bereich der beruflichen Weiterbildung,” *Zeitschrift für Arbeitsmarkforschung*, forthcoming.

Schräpler, J. P. (2004): “Respondent Behavior in Panel Studies. A Case Study for Income Nonresponse by Means of the German Socio-Economic Panel (SOEP),” *Sociological Methods & Research*, 33, 118–156.

Schupp, J., and G. G. Wagner (2002): “Maintenance of and innovation in long term panel studies: The case of the German Socio-Economic Panel (GSOEP),” *Allgemeines Statistisches Archiv*, 86, 163–175.

Skrondal, A., and S. Rabe-Hesketh (2004): *Generalized Latent Variable Modeling. Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall/CRC, Boca Raton.

——— (2005): *Multilevel and Longitudinal Modelling Using Stata*. Stata Press, College Station, Texas.

Sloane, P. J., and I. Theodossiu (1998): “Methodological and Econometric Issues in the Measurement of Low Pay and Earnings Mobility,” in *Low Pay and Earnings Mobility in Europe*, ed. by P. J. S. Asplund, Rita, and I. Theodossiu, pp. 3–12. Edward Elgar, Chelthenham.

Sousa-Poza, A. (2004): “Is the Swiss Labor Market Segmented? An Analysis Using Alternative Approaches,” *Labour*, 18, 131–161.

Statistisches Bundesamt (2001): *Statistisches Jahrbuch 2001. Für die Bundesrepublik Deutschland*. Statistisches Bundesamt, Wiesbaden.

——— (2006): *Statistisches Jahrbuch 2006. Für die Bundesrepublik Deutschland*. Statistisches Bundesamt, Wiesbaden.

Steiner, V. (2001): "Unemployment Persistence in the West German Labour Market: Duration Dependence or Sorting?," *Oxford Bulletin of Economics and Statistics*, 63, 91 – 113.

——— (2004): "Social Welfare Reform and the Low-Wage Labor Market in Germany: What Works and What Doesn't?," *Applied Economics Quarterly Supplement*, 50, 57–82.

Stevens, A. H. (1999): "Climbing out of poverty, Falling back in: Measuring the Persistence of Poverty over multiple Spells," *Journal of Human Resources*, 34, 557–588.

Stewart, J., and M. D. Dooley (1999): "The Duration of Spells On Welfare and Off Welfare among Lone Mothers in Ontario," *Canadian Public Policy*, 25 Supplement 1, 47–72.

Stewart, M. B. (2006): "The Inter-related Dynamics of Unemployment and Low-wage Employment," *Journal of Applied Econometrics*, forthcoming.

Stewart, M. B., and J. K. Swaffield (1999): "Low Pay Dynamics and Transition Probabilities," *Economica*, 66, 23–42.

Sueyoshi, G. T. (1992): "Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-Varying Covariats," *Journal of Econometrics*, 51, 25–28.

Train, K. (2000): "Halton Sequences for Mixed Logit," *Working Paper*, E00-278.

Train, K. (2003): *Discrete Choice Models using Simulation.* Cambridge University Press, Cambridge, UK.

UHLENDORFF, A. (2004): "Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer," *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 56, 279–303.

UHLENDORFF, A., AND K. F. ZIMMERMANN (2006): "Unemployment Dynamics among Migrants and Natives," *CEPR Discussion-Paper*, 5872.

VAN DEN BERG, G. J. (2001): "Duration Models: Specification, Identification, and Multiple Durations," in *Handbook of Econometrics, Vol.5*, ed. by J. J. Heckman, and E. Leamer, pp. 3381–3460. North-Holland, Amsterdam.

VOGES, J., AND G. ROHWER (1992): "Receiving Social Assistance in Germany: Risk and Duration," *Journal of European Social Policy*, 33, 175–191.

WACHTER, T. V., AND S. BENDER (2006): "In the Right Place at the Wrong Time - The Role of Firms and Luck in Young Worker's Careers," *American Economic Review*, 96, 1679–1705.

WILDE, J. (2003): "Why do Recipients of German Social Assistance Opt Out? An Empirical Investigation of Incentives with the Low Income Panel.," *Jahrbücher für Nationalökonomie und Statistik*, 223, 719–742.

WILDE, J., AND A. KUBIS (2005): "Non-take-up Behavior of Social Assistance in Germany," *Jahrbücher für Nationalökonomie und Statistik*, 225, 347–373.

WINKELMANN, R., AND K. F. ZIMMERMANN (1993): "Ageing, Migration, and labour mobility," in *Labor Markets in an Ageing Europe*, ed. by P. Johnson, and K. F. Zimmermann, pp. 255–282. Cambridge University Press, Cambridge.

WOOLDRIDGE, J. (2005): "Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20, 39–54.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, London.

Zimmermann, K. F. (2003): “Beschäftigungspotentiale im Niedriglohnsektor,” *Vierteljahreshefte zur Wirtschaftsforschung*, 72, 11–24.

——— (2005): “European Labour Mobility: Challenges and Potentials,” *De Economist*, 153, 425–450.

List of Tables

List of Figures

German summary

Übergänge auf dem Arbeitsmarkt: Arbeitslosigkeit, Transferbezug und der Niedriglohnsektor

Die Arbeitslosigkeit in Deutschland ist in den vergangenen Dekaden nahezu kontinuierlich angestiegen. Während die Arbeitslosensquote in Rezessionen zunimmt, geht sie in Zeiten des Aufschwungs nicht in gleichem Maße zurück. Anders gesagt ist die Arbeitlosigkeit nicht mehr vor allem ein zyklisches Phänomen, sondern struktureller Natur. Im Jahr 2005 waren ca. fünf Millionen Personen oder 13% der Erwerbspersonen als arbeitslos registriert.[1] Die Reduktion dieser hohen Arbeitslosenquote ist eine der dringensten wirtschaftspolitischen Herausforderungen.

Insbesondere die Gruppe der Geringqualifizierten hat ein hohes Arbeitslosigkeitsrisiko. Der Zusammenhang des Bildungsniveaus und der Wahrscheinlichkeit, arbeitslos zu sein, hat dabei in den letzten Jahren zugenommen. Zudem entwickeln sich seit den 1970er Jahren die Arbeitslosigkeitsraten der Einheimischen und der Migranten auseinander. 2005 war die durchschnittliche Arbeitslosenrate in der Gruppe der Ausländer ungefähr doppelt so hoch wie die innerhalb der Deutschen. Die Ursache hierfür kann in der unterschiedlichen Zusammensetzung beider Gruppen liegen. So ist bpsw. der Anteil der Geringqualifizierten unter den Migranten relativ hoch. Es ist aber auch möglich, dass unabhängig von beobachtbaren Eigenschaften einige Ethnien spezifische Nachteile auf dem Arbeitsmarkt aufweisen.

Es wird oft angeführt, dass die Etablierung eines Niedriglohnsektors für die Überwindung der hohen Arbeitslosigkeit unter den Geringqualifizierten notwendig sei und dass das aktuelle Steuer- und Transfersystem insbesondere im Niedriglohnbereich nur geringe Anreize zur Arbeitsaufnahme impliziert. In diesem Zusam-

[1] Die Arbeitslosenquote ist bezogen auf abhängige zivile Erwerbspersonen.

menhang wurden verschiedene Reformvorschläge wie die Reduktion der Sozialhilfe, die Einführung von Workfare oder von Kombi-Löhnen gemacht.

Die vorliegende Arbeit stellt einen Beitrag zur Debatte über die Determinanten individueller Beschäftigungsdynamiken dar. Hierbei geht es insbesondere um die Wahrscheinlichkeit, in die Arbeitslosigkeit und den Transferbezug einzutreten und diese Zustände wieder zu verlassen. Warum treten einige Haushalte aus der Sozialhilfe aus und andere nicht? Bleiben Migranten länger arbeitslos oder sind ihre Beschäftigungsverhältnisse kürzer im Vergleich zu Einheimischen? Existiert in Deutschland ein kausaler Zusammenhang zwischen niedrig entlohnter Beschäftigung und dem Risiko, arbeitslos zu werden? Diese Fragen werden in der vorliegenden Arbeit behandelt.

Alle empirischen Studien basieren auf den Daten des Sozio-ökonomischen Panels (SOEP), einer repräsentativen Längschnittstudie privater Haushalte in Deutschland. Die Daten beinhalten detaillierte Informationen zum Beschäftigungsstatus und dem Einkommen aller Haushaltsmitglieder sowie Informationen zum Migrationshintergrund. Der Längsschnittcharakter des SOEP erlaubt die Analyse inidivueller Arbeitsmarktdynamiken, die durch Verweildauern in bestimmten Zuständen sowie durch Übergangswahrscheinlichkeiten zwischen den Zuständen beschrieben werden.

Der erste Teil der Arbeit beschäftigt sich mit dem Einfluss der Sozialhilfehöhe auf die Abgangswahrscheinlichkeit aus der Sozialhilfe. Es wird häufig angeführt, dass die Höhe der Sozialhilfe für Arbeitskräfte mit geringer Produktivität nur geringe Anreize zur Arbeitsaufnahme impliziert. Um diesen Zusammenhang zu analysieren, ermittle ich den Einfluss des Verhältnisses zwischen geschätztem potentiellen Erwerbseinkommen und der Höhe des Sozialhilfeniveaus auf die Dauer des Sozialhilfebezugs. Das potentielle Netto-Erwerbseinkommen wird unter Berücksichtigung des Steuersystems auf Grundlage einer Lohnschätzung mit Selektionskorrektur ermittelt. Es werden Übergangsratenmodelle in diskreter Zeit mit konkurrierenden Risiken und unter Berücksichtigung unbeobachteter Heterogenität geschätzt. Die Ergebnisse zeigen, dass das Verhältnis einen positiven Effekt auf die Wahrscheinlichkeit, die Sozialhilfe zu verlassen, hat. Dieser Effekt ist insbesondere für Haushalte relevant, deren potentielles Einkommen die Sozialhilfehöhe übersteigt. Die Ergebnisse unterscheiden sich von den Ergebnissen vorheriger Studien, die die Verweildauer in Sozialhilfe in Deutschland untersucht haben, entsprechen aber den Ergebnissen der internationalen Literatur. Die Ergebnisse meiner Studie resultieren zum einen aus der simultanen Berücksichtigung beider Einkommensquellen, der Sozialhilfehöhe und des potentiellen Erwerbseinkommens. Denn nur durch die Berücksichtigung beider Einkommensquellen lässt sich die mit dem Übergangsprozess zwischen beiden Zuständen verbundene Anreizstruktur adäquat abbilden. Zum anderen ist eine Unterscheidung der konkurrierenden Risiken "Beschäftigung" und "alternative Übergänge" notwendig, denn der Einfluss des potentiellen Erwerbseinkommens ist nur bei Übergängen in Beschäftigung relevant.

Im zweiten Teil der Arbeit geht es um die Arbeitsmarktdynamiken von Migranten. Migranten sind zu einem größeren Anteil arbeitslos als Einheimische. Diese höhere Arbeitslosenquote kann aus einem höheren Risiko, arbeitslos zu werden, resultieren, also aus kürzeren Beschäftigungsverhältnissen bzw. häufigeren Arbeitslosigkeitsphasen. Daneben kann der Unterschied auch aufgrund von längeren Arbeitslosigkeitsphasen, also einer geringeren Abgangswahrscheinlichkeit aus der Arbeitslosigkeit auftreten. Ich untersuche beide möglichen Quellen höherer Arbeitslosigkeitsraten, die Arbeitslosigkeitsdauer und die Stabilität nachfolgender Beschäftigungsverhältnisse. Beide Prozesse werden von beobachtbaren und unbeobachtbaren Eigenschaften beeinflusst, wobei es realistisch erscheint, dass die unbeobachteten Eigenschaften, die die beiden Prozesse beeinflussen, nicht unabhängig voneinander sind. Daher schätze ich beide Prozesse simultan und erlaube für eine Korrelation zwischen den unbeobachteten Faktoren. Die Ergebnisse zeigen, dass Migranten mit gleichen beobachtbaren und unbeobachtbaren Eigenschaften längere Zeit brauchen, um eine neue Beschäftigung zu finden. Diese neuen Beschäftigungen sind dann allerdings genauso stabil wie diejenigen der Einheimischen. Die Wahrscheinlichkeit, die Arbeitslosigkeit zu verlassen, variiert stark zwischen den Ethnien, wobei Türken der ersten und zweiten Generation als die Gruppe mit den größten Problemen identifiziert werden.

Im dritten Teil der Arbeit untersuche ich Übergangsprozesse zwischen niedrig entlohnten Beschäftigungsverhältnissen und Arbeitslosigkeit. Geringqualifizierte Personen haben ein relativ hohes Arbeitslosigkeitsrisiko. In diesem Zusammenhang werden Niedriglohnbeschäftigungen kontrovers diskutiert. Auf der einen Seite wird argumentiert, dass steigende Beschäftigungszahlen im Niedriglohnsektor eine Lösung der hohen Arbeitslosenzahlen unter den Geringqualifizierten darstellen kann. Auf der anderen Seite werden niedrig entlohnte Beschäftigungen oft mit unstabilen Erwerbsverläufen und einem hohen Arbeitslosigkeitsrisiko assoziiert. In diesem Zusammenhang ist es wichtig zu wissen, inwieweit Niedriglohnbeschäftigungen von kurzer Dauer und Zwischenschritte zu besser bezahlten Beschäftigungsverhältnissen sind oder ob in Deutschland ein Kreislauf zwischen Niedriglohn-Jobs und Arbeitslosigkeit beobachtet werden kann. Ich untersuche die Dynamiken von Arbeitslosigkeit und Niedriglohnbeschäftigung für Männer in Westdeutschland. Der Fokus liegt auf dem Ausmaß der tatsächlichen Pfadabhängigkeit in diesen Zuständen. Ich schätze dynamische multinomiale Logit-Modell mit unbeobachteter Heterogenität. Hierbei werden sowohl die mögliche Endogenität des zuerst beobachteten Arbeitsmarktzustandes ("initial condition") als auch die mögliche Endogenität des Panelausfalls ("panel attrition") berücksichtigt. Übereinstimmend mit Studien für andere Länder findet diese erste Studie für Deutschland eine Pfadabhängigkeit in niedrig entlohnten Beschäftigungen. Zusätzlich bestätigen die Resultate vorherige Ergebnisse zur Pfadabhängigkeit in Arbeitslosigkeit. Darüber hinaus finde ich Evidenz für eine gegenseitige Abhängigkeit der beiden betrachteten Zustände: Gering bezahlte Beschäftigung geht mit einem höheren Arbeitslosigkeitsrisiko einher und umgekehrt. Allerdings führen Niedriglohn-Jobs im Vergleich zu Nicht-Arbeit zu einer höheren Beschäftigungswahrscheinlichkeit und mit einer höheren Wahrscheinlichkeit zu besser bezahlter Beschäftigung. In Bezug auf zukünftige Beschäftigungsaussichten ist es demnach besser, gering bezahlt zu sein als nicht beschäftigt zu sein.

Welche wirtschaftspolitischen Schlussfolgerungen lassen sich aus den Ergebnissen dieser Analysen ziehen? Für die Sozialhilfe zeigt sich, dass höhere Transferbezüge zu längerem Sozialhilfebezug führen. Eine Reduktion des Sozialhilfeniveaus wäre eine Möglichkeit, das Anreizproblem im Niedriglohnbereich zu lösen. Allerdings beläuft sich die Höhe der Sozialhilfe auf die Höhe des Existenzminimums und eine generelle Absenkung dürfte politisch nur schwer durchsetzbar sein. Eine Absenkung des Transferniveaus ist aber nicht die einzige Möglichkeit, das Anreizproblem der Sozialhilfe zu lösen. Eine weitere Möglichkeit stellt bspw. die Einführung von Workfare dar. Bei der Untersuchung der Arbeitsmarktsituation der Migranten werden die türkischen Arbeitslosen als Hauptproblemgruppe identifiziert. Daher sollte die Politik sich auf die Prozesse der Arbeitssuche dieser Gruppe konzentrieren, um die Nachteile der Migranten auf dem Arbeitsmarkt zu bekämpfen. Die Ergebnisse der Analyse der Nideriglohndynamiken zeigen, dass niedrig entlohnte Beschäftigungen im Vergleich zur Nichtbeschäftigung zu besseren Beschäftigungsaussichten führen. Daher sollte die Politik versuchen, den Niedriglohnsektor auszubauen.

Publications

Articles in Refereed Journals

Die Wirkungen der Hartz-Reform im Bereich der beruflichen Weiterbildung, *Zeitschrift für Arbeitsmarktforschung*, forthcoming (with Hilmar Schneider)

Estimation of multinomial logit models with unobserved heterogeneity using maximum simulated likelihood, *Stata Journal*, 6 (2006), 229-245 (with Peter Haan)

Transitions from Welfare to Employment: Does the Ratio between Labor Income and Social Assistance matter?, *Schmollers Jahrbuch*, 125 (2005), 51-61 (with Hilmar Schneider)

Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer, *Kölner Zeitschrift für Soziologie und Sozialpsychologie,* 56 (2004), 279-303

Work in Progress/Discussion Papers:

From Low Pay to No Pay and Back Again? A Multi-State Model of Low Pay Dynamics, 2006, DIW Discussion Paper No. 648/ IZA Discussion Paper No. 2482

Unemployment Dynamics among Migrants and Natives, 2006, DIW Discussion Paper No. 617/ IZA Discussion Paper No. 2299/ CEPR Discussion Paper No. 5872 (with Klaus F. Zimmermann)

Intertemporal Labor Supply and Involuntary Unemployment in a Rationed Labor Market, mimeo (with Peter Haan)

Self-Employment Dynamics, Cross-Mobility Patterns and True State Dependence, mimeo (with Marco Caliendo)

Reforming Publicly Financed Training in Germany: Quality and Selection Effect, mimeo (with Ulf Rinne and Zhong Zhao)

Too Bad to Benefit? Effect Heterogeneity of Publicly Financed Training in Germany, mimeo (with Ulf Rinne and Marc Schneider)

Evaluating Continuous Training Measures Using the Generalized Propensity Score, mimeo (with Jochen Kluve, Hilmar Schneider and Zhong Zhao)

Further Publications:

Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission - Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen, IZA Research Report, Bonn, 2006 (with H. Schneider, K. Brenke, D. Hess, L. Kaiser and J. Steinwede)

„Arbeitslosigkeit“, DIW@ School, Nr. 1/2005 (with Peter Haan and Katharina Wrohlich)

Lohnkosten im internationalen Vergleich, DIW Wochenbericht 14/2004, 161-169 (with Tilman Brück and Malte Woweries)