An Empirical Model of Life-Cycle Earnings and Mobility Dynamics *

Florian Hoffmann

Department of Economics

University of Toronto

JOB MARKET PAPER
December 2009

Abstract

Conventional estimates of empirical human capital investment models of post-graduation career dynamics suggest that pre-labor market skills are the predominant source of life-cycle earnings inequality. In this paper I test if this conclusion is significantly altered when a proto-typical dynamic Roy model of life-cycle income dynamics and vertical occupational mobility is enriched with a number of potentially important sources of career heterogeneity, such as match heterogeneity, search frictions, and permanent shocks to skills. I estimate the parameters of the resulting structural model using a unique administrative Panel Data Set which follows a large sample of employees with identical educational attainments from the time of their labor market entry until twenty-three years into their careers. I find that a large fraction of life-cycle income inequality is driven by match heterogeneity among workers with the same observable and unobservable credentials. Differences in comparative advantages, though quantitatively important as well, have a much smaller impact than what has been found in research that relies on estimates from more restrictive dynamic Roy models. Thus, compared to the conclusions drawn from models which do not control for unobserved sources of career heterogeneity that accumulate over a life-cycle, my results suggest that policies targeting pre-labor market skill accumulation are likely to be less effective, and active labor market policies are likely to be more effective in fostering career progression.

*This study uses the weakly anonymous IAB Employment Sample. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and remote data access. I thank Philip Oreopoulos, Shouyong Shi and Victor Aguirregabiria for their help and continuous support throughout this project. I also thank Joe Altonji, Gueorgui Kambourov, Fabian Lange, Diego Restuccia, Aloysius Siow and seminar participants at the CEPA seminar at the University of Toronto, the RMM 2009 conference and the Yale Labor-Public Finance Lunch Seminar for helpful suggestions, and Benedikt Hartmann, Daniela Hochfellner, Peter Jacobebbinghaus and other staff members at the FDZ-IAB for their hospitality. Financial support from Philip Oreopoulos, Shouyong Shi's Canada Research Chair Fund, the CLSRN fellowship, the University of Toronto Travel Grant and the IAB is gratefully acknowledged. All remaining errors are mine.

1. Introduction

It is well known that earnings and wages vary substantially over an individual's life-cycle. Workers who enter the labor market with the same observational credentials can experience very different career trajectories and end up at very different positions in the earnings distribution. Whether this career heterogeneity is predominantly driven by skills acquired prior to labor market entry, such as educational attainment, parental investments or innate abilities, or by decisions made and shocks accumulated over the working life, remains highly controversial. The design of policies that effectively promote success in the labor market relies on an empirically well-founded answer to this question. Advocates of the view that career outcomes are mainly driven by abilities acquired when young stress the importance of policies that foster early childhood education, increase and homogenize the quality of primary and secondary education, and support parental investment throughout a youth's development.[1] In contrast, researchers who find evidence for a large role of luck and exogenous shocks that occur over the career emphasize the need to implement active labor market policies and social insurance programs.[2]

In this paper I develop and estimate a structural model of life-cycle earnings and mobility dynamics in which individuals have to choose between occupation groups of different quality and unemployment in each period of their career. The model structure is built on a dynamic Roy-model of occupational mobility with endogenous human capital accumulation as considered in Keane and Wolpin's (1997) seminal study of career dynamics of young men. Possibly due to computational constraints and data quality, Keane and Wolpin and the majority of subsequent work that estimates Dynamic Discrete Choice models of post-graduation career dynamics abstract from many important determinants of earnings and mobility dynamics, such as permanent shocks to skills, match heterogeneity and search for better job opportunities. As I argue and demonstrate in my analysis, this potentially leads to an upward bias in the estimated role of pre-labor market skills for career development.

[1]See for example Restuccia and Urrutia (2004); Heckman, Stixrud and Urzua (2006), Cunha (2007), Cunha and Heckman (2007, 2008), and Caucutt and Lochner (2008).

[2]For work exploring the role of social insurance in fostering career progression see e.g. Huggett and Ventura (1999), Low, Meghir and Pistaferri (2008) and Blundell, Pistaferri, Preston (2009).

In particular, differences in post-graduation outcomes that are due to factors omitted from their analysis are, at least to some extent, interpreted as outcomes from differences in pre-labor market skills. For example, two individuals who enter the labor market with identical sets of observable and unobservable skills and initially work in the same occupation may experience very different career outcomes in the presence of labor market frictions because one of them is "lucky" and finds a particularly good career match, while the other is "unlucky" and does not have a favorable outcome to job search.[3] Consequently, the former experiences a discrete and permanent jump of wages at the time of a job change that, when not controlled for in the structural model, will be interpreted as differences in comparative advantages established prior to labor market entry.[4] Similarly, permanent unobserved shocks to skills, such as an accident or a job promotion or demotion within occupational classes, generate permanent income inequality among otherwise identical individuals and, when excluded from the model, will be erroneously subsumed in the estimates of pre-labor market skills.

Thus, to reach a comprehensive model of post-graduation earnings and mobility dynamics, I integrate a flexible model of residual earnings dynamics into a Dynamic Discrete Choice Model of career progression, human capital accumulation, and partial equilibrium search. As in Keane and Wolpin's model, pre-labor market skills, entering the model as unobserved permanent differences in comparative advantages, determine the quality of jobs individuals initially sort into. Workers endogenously accumulate human capital as their career progresses. I augment the Keane-Wolpin model by introducing match heterogeneity, search, and exogeneous permanent shocks to skills. Match effects remain constant for the duration of a job.[5] This generates a jump process of residual wages, with jumps occuring at the time of mobility. Consequently, the parameter describing the distribution of matches is identified from the systematic residual variation at the time of an occupational change.

[3] See Burdett and Mortensen (1998) for an equilibrium search model in which firms offer different wages to identical workers in equilibrium.

[4] Topel and Ward's (1992) finding that one third of wage growth in US administrative data is explained by job mobility suggests this bias to be substantial.

[5] Jovanovich's model (1979) is seminal in the literature on search for matches. Empirical applications following the structure of his models are Miller (1984), Neal (1999) and Pavan (2009).

Unlike most applications of discrete choice modeling in Labor Economics in which only choice probabilities are considered - usually using linear probability models or Probit - I derive the joint likelihood of the entire individual-specific history of observed choices, wages and predetermined variables. The likelihood does not have a closed form. I thus refer to Maximum Simulated Likelihood Estimation. Precise estimation of the model parameters requires data with long individual-level labor market histories, a large cross-section and little measurement error. A high quality administrative Panel Data Set from Germany, that follows 56,000 employees from the time of labor market entry until twenty-three years into their careers, satisfies all these criteria and provides me with sufficient statistical power. Furthermore, since unemployment and social insurance benefits collected by a worker are recorded in the data, I can consider a flexible parameterization of the social insurance system. Focusing on a sample of workers with identical educational attainments, apprenticeship experiences and ages at labor market entry allows me to abstract from the endogeneity of initial conditions. I show that my model specification significanly improves upon the Keane-Wolpin framework in terms of its match to multiple dimensions of the data. Wage-profiles, life-cycle profiles of employment shares observed in each choice, and mobility rates across alternatives, are matched very well.

I begin my empirical analysis with estimating a version of the Keane-Wolpin model that is nested within my full structural model. Although I use different data, I reach at an almost identical estimated role of pre-labor market skills for life-cycle earnings inequality. In particular, 91 percent of the variation of life-cycle earnings is explained by type heterogeneity, compared to 90 percent as found in Keane and Wolpin for individuals in the NLSY. However, once I introduce permanent shocks to skills, match heterogeneity and search, my conclusions change dramatically. The estimated role of pre-labor market skills decreases from 91 to only 41 percent. Furthermore, using counterfactual exercises I find that excluding match heterogeneity from the full model decreases the standard deviation of life-cycle earnings by 27 percent. Thus, a large fraction of life-cycle income inequality usually interpreted as the outcome of type heterogeneity is in fact rooted in match heterogeneity among workers of the same skill type. Employees who are initially "endowed" with the same abilities and enter the labor market with identical observable and unobservable credentials experience very different career trajectories because they have different outcomes to job search and are hit by different permanent shocks. When not controlling for these systematic earnings changes that take place over a career, they will be included in the estimates of skills that are initially carried into the labor market.

I provide evidence that the importance of initial match quality does not reflect unobserved permanent ability differences not controlled for in the model. Although the latter cannot be separately identified within my model, its predictions on life-cycle career dynamics can be contrasted with the predictions of match heterogeneity. In particular, if match quality was driven by uncontrolled ability differences, a large mass of workers with very low initial earnings would never catch up, generating too large a life-cycle earnings inequality and earnings immobility. I also establish that controlling for match heterogeneity leaves very little within-match residual earnings variation. This, as I argue, rules out match heterogeneity within occupational hierarchies to be primarily driven by inter-firm mobility. Instead, my results suggest that firm mobility is most important if it is associated with an occupational up-or downgrade.

In contrast to single-equation models of earnings dynamics, my model features several dimensions of state-dependence in the sense that individuals observed in different employment states and occupations might differ by the exogenous risk they are subjected to, how their skills are valued, and how they adjust their behavior to different kinds of shocks. It is thus well suited for the study of income mobility as defined by the probability that an individual in the p-th quantile of experience-specific earnings distributions is observed in the q-th quintile some time later. Consequently I apply the same counterfactual experiments conducted for the study of life-cycle income inequality to the study of income mobility. I find that the exclusion of type- or match heterogeneity has quantitatively comparable impacts on the outcome of interest. Both, in a counterfactual world without type or match heterogeneity, income mobility would be significantly higher. However, while in the former case extreme transitions become more likely, smooth transitions are more frequent in the latter case.

I conclude that policies targeting pre-labor market skill accumulation are likely to be less effective, and active labor market policies are likely to be more effective in fostering career development than commonly accepted. Both policies, when implemented efficiently, would help avoiding poverty traps and increase earnings. These conclusions only apply for ability differences within educational groups and does not speak to the literature estimating returns to education.

My work contributes to a growing literature that formulates and estimates decision-theoretic models of labor market mobility and income dynamics. Topel and Ward (1992) use administrative data from the US to estimate empirical hazard functions of job mobility that condition on wages. They find that at least one-third of early-career wage growth is explained by wage gains experienced at job changes. This suggests that estimating models of the joint dynamics of earnings and mobility is crucial to reach credible estimates of the sources of life-cycle career progression. Starting with the seminal work by Keane and Wolpin (1997), dynamic models with a finite set of alternatives have gained increasing popularity in the analysis of post-graduation career dynamics and policy analysis.[6] Two recent applications, Adda, Dustmann, Meghir and Robin (2009) and Adda, Costa Dias, Meghir and Sianesi (2009), estimate dynamic discrete choice models with permanent exogenous skill shocks and apply them to policy evaluation. In contrast to my model, the discrete choices in their works are different policy regimes, such as vocational training versus labor market experience, or certain social programs, while post-graduation dynamics are summarized in one wage equation.[7]

The rich specification of the institutional environment and of unobserved heterogeneity

[6] Starting with Rust's (1987) study there is a rich literature in Industrial Organization addressing the estimation of dynamic discrete choice models. For a recent survey refer to Aguirregabiria and Mira (2009). In labor and public economics, this framework is still relatively unused. Sullivan (2009) is a recent extension of Keane and Wolpin's seminal study.

[7] There are a number of subtle, but important differences between these two studies and the model considered in my work. First, in their framework, once individuals have sorted into one of the policy regimes, such as vocational training, no discrete choice has to be made anymore. Although there is mobility across firms, the wage and career progress after the initial sorting is summarized by one equation. Second, they assume that match effects follow a random walk that is initialized each time a worker changes employers or becomes unemployed. In contrast, in my model match effects and the random walk are two distinct objects following separate dynamic laws of motion. This is crucial for the focus of my paper.

usually considered in Dynamic Discrete Choice and Dynamic Probit Models comes at the cost of keeping the model in a partial equilibrium framework. Equilibrium search models, in which the distribution of match effects is endogenous, are an interesting alternative.[8] However, to be computationally tractable, these models can admit neither serially correlated errors, nor non-stationarity and are thus not well suited for quantifying the relative importance of pre-labor market skills and of different sources of risk for life-cycle earnings inequality.[9]

By merging a Dynamic Roy model with a partial equilibrium model of search for better matches and a flexible residual variance components model, my work also contains some technical innovations. There is considerable interest in the estimation of Dynamic Discrete Choice Models and Dynamic Probit Models in areas such as Labor and Public Economics, Industrial Organization and Computational Economics.[10] Given the non-linearity of the models, they generally do not admit closed-form solutions. Consequently, the researcher needs to rely on simulation-based estimation, and modeling choices are limited by data quality and computational tractability. Most applications allow for transitory choice-specific shocks only, thus imposing arguably strict assumptions on the error process and consequently on the theoretical choice functions. My work is the first to integrate serially correlated shocks to general skills and dynamically evolving match heterogeneity into a discrete choice framework.

I also contribute to a large and still growing literature that attempts to estimate flexible statistical variance components models of post-graduation earnings dynamics.[11] Different

[8]Papers estimating equilibrium search models are e.g. Eckstein and Wolpin (1990), Bontemps, Robin and van den Berg (1999), Postel-Vinay and Robin (2002), Cahuc, Postel-Vinay and Robin (2006) and Lise (2007).

[9]Some progress has been made with respect to the non-stationarity assumption. See for example Moscarini and Postel-Vinay (2009), Shi (2009) and Shi and Menzio (2008, 2009). Postel-Vinay and Turon (2009) and Bagger, Fontaine, Postel-Vinay and Robin (2006) specify equilibrium search models with human capital accumulation. In all these works, unobserved heterogeneity is still very restrictive.

[10]For work about the numerical implementation of Dynamic Discrete Choice models, see e.g. Hotz and Miller (1993), Hotz, Miller, Sanders and Smith (1994), Aguirregabiria and Mira (2002, 2007), Imai, Jain and Ching (2006) and Norets (2009).

[11]Well cited studies estimating models of earnings dynamics are Hause (1980), MaCurdy (1982), Gottschalk and Moffitt (1994, 2002), Baker (1997), Haider (2002), Baker and Solon (2003), and Guvenen (2009). Hoffmann (2009) estimates a very flexible non-stationary model with the same data used

variance components motivate very different policy recommendations. On the one hand, since very persistent shocks cannot be smoothed by savings or other private insurance mechanisms, it can be socially efficient to provide social insurance against them. On the other hand, the welfare effect of transitory shocks is too small to call for publicly financed insurance. Inherent in this line of research, conducted in a variety of areas such as Labor Economics, Public Economics and Macroeconomics, is the assumption that unobserved shocks are exogenous.[12] Abowd and Card (1989) are among the first to recognize and test for the potential endogeneity of persistent shocks by incorporating a dynamic variance components model into a standard life-cycle labor supply model. Two very recent studies - Altonji, Smith and Vidangos (2009) and Low, Meghir and Pistaferri (2009) - correct flexible statistical wage processes for mobility across employment states and across firms using a system of selection equations.[13] Overall, the conclusion is that a considerable fraction of dynamic earnings shocks usually interpreted as exogenous risk is indeed endogenous and acted upon. Low et. al (2009) demonstrate the importance of this finding for the design of social insurance programs.

A strand of this literature estimates firm and worker fixed effects using matched employer-employee data and quantifies the importance of firm and worker heterogeneity. Abowd, Kramarz and Margolis (1999) and research built on their empirical model generally find that firm heterogeneity is important, pointing towards a strong relationship between worker mobility and earnings dynamics. Recent papers by de Melo (2009) and Lentz (2009) address similar issues in structural equilibrium search models. Kambourov and Manovskii (2009) identify horizontal occupational mobility as a catalyst of wage inequality.

Comparable to these works, I use the empirical mobility decisions of individuals on the Micro-level to extract information about the sources of residual earnings dynamics.

herein.

[12]Early examples of research investigating behavioral and welfare implications of the persistence of individual earnings shocks are Hall and Mishkin (1982) and Quah (1990). Recent examples include Gourinchas and Parker (2001, 2002), Haider (2001), Heathcote et al (2005, 2008), Krueger and Perri (2006), Huggett et al (2006, 2007), Storesletten et al (2004 a, b), Kaplan (2007) and Guvenen (2007). Examples of studies that explore the relationship between individual wage and consuption processes are Blundell, Pistaferri and Preston (2009) and Guvenen and Smith (2009).

[13]An earlier study in this literature is Altonji, Martins and Siow (2002).

However, rather than maximizing the flexibility of the statistical model of the dynamics of wages and other labor market variables, I embed a dynamic variance components model into a behavioral framework of occupational choice. Latent wage and earnings equations are replaced by structural wage equations, and individuals need to solve an optimization problem. This makes the model well suited for long-run policy analysis, but also for incorporating it into a unified framework of earnings, mobility and consumption dynamics.

The rest of the paper is organized as follows: In Section 2 I describe the data and the heuristic used to construct the occupational classes, and I establish a number of stylized facts about career progression. In the sub-sequent section I lay out the model structure and discuss its main features. Estimation and identification is discussed in section 4, followed by a section presenting the estimation results and the match of the model. Section 6 quantifies the role of different potential sources of life-cycle inequality and income mobility by relying on counterfactual experiments. I finish with a conclusion and a discussion of model extensions and further applications.

2. Data and Descriptive Statistics

2.1. Data Description

I use the confidential version of the IABS, a 2%-extract from German administrative social security records.[14] For the purpose of this study, using these data instead of publicly available Panel Data has at least 5 advantages: First, I can generate unusually long series of wage observations for the same individuals - in my final sample I observe up to 23 wage and unemployment benefit records for the same worker. Second, earnings histories are observed from the time of labor market entry, considerably simplifying the treatment of initial conditions. Third, the IABS provides a well-defined education variable. This

[14] These data are collected by the "Institut fuer Arbeits-und Berufsforschung" (IAB) (Institute for Employment Research) at the German Federal Employment Agency.

enables me to perform separate analyses for each education group. At this point, for reasons explained below, I focus on the largest education group in the German Labor market. Fourth, given that wage records are provided by firms under the threat of being severly punished for misreporting, measurement error is arguably minimized. Fifth, by observing the exact amount of unemployment benefits and social insurance collected by an individual I can consider a sufficiently rich and realistic latent structural wage equation for this choice.

General Description of the IABS The IABS is a 2%-extract from German administrative social security records for the years 1975 to 2004. Once an individual is drawn, it is followed for the rest of the sample period. The IABS is representative of the population of workers covered by the social security system. Excluding self-employed and civil servants, this amounts to approximately 80% of the German workforce. In order to keep the sample representative, a new random sample of labor market entrants is added each year.

Work spells and unemployment spells are recorded with exact start and end dates. A spell ends for different reasons, such as a change in employment status, a change in employer or occupation, or a change in whether the worker is working full- or part-time. If no such change occurs, a firm has to report one spell per year for each of its workers. The data report average daily wages per spell. For this reason I will restrict the sample to those working full-time or being unemployed. To keep the sample tractable and computation feasible I aggregate the records up to the annual level. Given that mobility rates strongly peak between December and January, with relatively low mobility rates during the rest of the year, the aggregation bias can be expected to be low.

Transitions in and out of self-employment and in and out of the labor force are potentially affecting my sample. Since I cannot observe these types of spells I only keep individuals whose labor market history starts with the year of labor market entry and ends with the most recent sample year. I therefore have a balanced panel within each cohort.

Sample Restrictions I restrict the sample to male full-time workers observed from the time of labor market entry. Starting in 1990, as a consequence of the German Unification, the sample also adds records from Eastern Germany. I focus on workers whose whole history of spells is recorded in Western Germany. This minimizes the possibility of earnings dynamics being driven by institutional changes related to the Unification.

The age at labor market entry varies considerably. Some individuals enter the sample when they are quite old, possibly because they change from self-employment or public employment into the status as private sector employee, or change from non-participation to labor market participation. This decision is potentially endogenous. To avoid initial conditions problems I construct a group of "typical" labor market entrants: In the first step I compute empirical mass points of age at labor market entry for each education group. Subsequently I drop individuals who entered after or at least two years before this age.

A further initial conditions problem is introduced from the endogeneity of an employee's educational attainment. Modelling this decision is potentially very difficult given the limited variation of educational policies across German provinces. Estimating choice rules thus relies on strong exclusion restrictions.[15] Since I focus on post-graduation earnings and mobility dynamics I do not model the education choice. Instead I keep the largest education group only which constitutes around 80% of the total IABS sample. Unlike in the sample of the highest educated, this group's fraction of top coded wages is low and comparable to publicly available US Panel Data Sets. Another rationale for focussing on one education group relates to the incidence of mobility. Studies dividing the sample into blue- and white collar occupations without restricting it to one education group only artificially deflate mobility rates because highly educated workers mostly select into the latter occupations, while the rest selects into the former.

The education variable provided by the IABS has 6 categories, ranging from "no degree at all" to "university degree". This variable is not necessarily constant over an individuals'

[15] Adda et al. (2009) estimate the returns to vocational training in the German labor market. Their exclusion restriction is motivated by a law-of-one-price which postulates that business cycle variation influences province-specific availability of trainee positions, but not relative wages across provinces. Given the different focus of my paper I do not consider such a model of educational choice.

labor market history. If for example an individual changes training status, education increases from "no degree at all" to "vocational degree". In some cases, the education variable decreases over time, mostly when there is a change in the firm of employment. To generate a consistent time series of education, I first use a refined version of the algorithm described in Fitzenberger et al. (2007). Subsequently, I generate a variable recording an individuals' highest level of education. Finally, I aggregate this variable up to three categories, "no degree at all", "highschool and/or vocational degree" and "post-secondary degree". I only keep the middle group, with the "typical" employee entering the labor market at age of 23. I thus drop individuals who entered past this age or before the age of 22. Only cohorts with both, sufficiently large cross-sectional sample sizes and labor market histories of at least 5 years length, are kept. This leaves me with the 18 cohorts born between 1955 and 1972. The longest labor market history in the data thus starts in 1978 and ends in 2004. To keep the number of observations for each experience group in the sample large enough I drop those observations with potential labor market experience above 22 years.[16]

The remaining sample has 851,375 observations. Computation times are still too long.[17] I thus draw a 10% random sample of individuals and keep the entire labor market histories of the remaining employees. Sample sizes by potential experience are shown in table 1 for both, the full sample and the 10% sub-sample. Initially, there are 55,677 (5,592) individuals in the sample. This number remains constant over the first 5 years by construction of the sample and then monotonically decreases to a sample size of 8,632 (873) after 23 years of labor market experience.

2.2. Definition of Occupational Classes

Vertical occupational mobility is frequently cited as an important part of an individuals' career progression. In contrast to horizontal occupational mobility, occupational upgrades, such as a move from being a bank teller to a branch manager, represent discrete changes in

[16]It is important to note that experience in the first year is equal to zero. Thus, a value of 22 for experience is associated with a labor market history of 23 years length.

[17]Adda et al (2009a and 2009b) face the same problem.

career trajectories, often seen as mirroring "success". Unlike most administrative data sets the IABS provides accurate 3-digit occupational codes. I assign each of the 338 occupations in the sample to one of three occupational classes, refered to as "bad", "mediocre" and "good".[18] The assignment heuristic only utilizes variables not entering the model as exogenous or endogenous variables. Therefore, average wages or wage growth, both of which are endogenous in the model, cannot be used to define occupational categories. Instead I calculate the fraction of employees with a post-secondary education in each of the 338 occupations from a sample including post-graduate employment spells for any educational group. The one third of occupations with the lowest proportions of highly educated are labelled as "bad" and the one third of occupations with the highest proportions of highly educated are labelled as "good". The rest are labelled as "mediocre". Mechanically dividing the support of the distribution of occupation-level proportions of highly educated workers serves to keep sample sizes for each occupation group relatively large. "Good" occupations are mainly composed of white-collar jobs, while "bad" occupations contain a large fraction of low-skill blue collar occupations.

Figure 1 plots the fraction of highly educated against the average wage on the occupational level. There is a pronounced positive relationship between the two variables, with a large group of occupations without university graduates and very low wages and a cluster of occupations with above-average wages and a very high fraction of university graduates.[19] The former group is entirely covered by "bad" occupations, while the latter is entirely contained within the "good" occupations. In between these two groups of extremely low or high fraction of university graduates there is a long and flat profile where the fraction of university graduates and wages both increase. Some ot these occupations are still covered by the "bad" group, but the majority falls into the "mediocre" group. Thus, without relying on wages to define occupational groups, the heuristic automatically defines low versus highly paid occupations. I repeat the same exercise, but only keeping wage records of the estimation sample. This rules out that the relationship between average

[18] This terminology is quite strong but serves to clarify that my heuristic generates a hierarchy among the occupations.

[19] This relationship is statistically significant.

wages and the fraction of highly educated is driven by positive returns to education. Figure 2 shows that the qualitative features from Figure 1 are kept, but with considerably ore noise around the trend line. Figure 3 plots the full non-parametric earnings distributions in the working sample for the three occupational classes. While earnings distributions of "Bad" and "Mediocre" occupations are very similar, with a slightly higher mass of high wages in the latter, "Good" occupations are associated with a much higher fraction of highly paid individuals.

2.3. Stylized Facts

In this section I describe the main stylized facts regarding life-cycle wage and mobility dynamics. This serves the purpose to establish a number of interesting empirical regularities associated with earnings dynamics, occupational mobility, and the relationship between the two. Furthermore, the descriptive statistics are used to investigate the empirical match of the model once the estimation has been conducted. For the purpose of comparison I compute all descriptive statistics for the full working sample and its 10-percent sub-sample relied upon in the estimation.

In the following, a transition from alternative $j(t)$ in period t to alternative $j(t+1)$ in period $t+1$ is defined by the conditional probability $\frac{\Pr(j(t) \text{ to } j(t+1))}{\Pr(j(t))}$, and a discrete wage change is defined to be a wage increase or decrease by more than 10 percent of the standard deviation of wages. Figures 4 to 11 together with Tables 2 and 3 show the following patterns:

1. Real Wages and their standard deviations increase monotoneously over the life-cycle, the former in a concave, the latter in a linear manner (figure 4). Residuals from regressions of log-wages on cohort fixed effects, a polynomial in general experience and occupation-specific tenure, a time trend and the unemployment rate decrease over the first five years of a career and then start to increase linearly (figure 5).

2. There is a large re-allocation of workers from "Bad" and "Mediocre" occupations to "Good" occupations over the life-cycle as reflected by a significant decrease of the

employment shares in the two former occupational classes and a strong increase in the latter (figure 6).

3. Transition rates from employment into unemployment and vice versa are initially quite high and decrease over the life-cycle. In both cases, these rates are very low for "Good" occupations compared to the other two occupation groups (figures 7 and 8).

4. Upward mobility is most pronounced for transitions from bad to mediocre occupations, followed by transitions from mediocre to good occupations. There is only very little mobility from bad to good occupations (figure 9).

5. In the other direction, transition rates from good to mediocre and from mediocre to bad occupations are frequent, but decreasing over the life-cycle. Mobility from good to bad occupations is almost non-existent (figure 10).

6. The association between occupational upgrading and discrete wage increases is stronger than the association between occupational downgrading and wage increases (table 2). Both are frequent.

7. The association between occupational upgrading and wage decreases is weaker than the association between occupational downgrading and wage decreases (table 2). Both are frequent, but less so than wage increases.

8. The non-parametric distribution of life-cycle earnings, as defined as the sum of earnings net of cohort effects, over the first 23 years of individual careers is right-skewed (figures 11 and 12).

9. Earnings mobility - the probability that a worker with income in the p-th percentile of experience-specific earnings distributions receives income in the q-th quintile one year later - is low. In particular, the probability that an individual remains in the same position within the earnings distribution, net of experience effects, ranges from 61 percent in the middle of the distribution to 82 percent at the upper end of the

distribution. Earnings mobility is the lowest at the tails of the distributions: Individuals are highly likely to remain poor or rich within a year. Five- and ten-year transition matrices of earnings exhibit a much lower degree of persistence, although it is still high.

The sub-sample replicates the profiles computed from the full sample well, but with considerably more "noise", especially at higher values of experience and for mobility profiles associated with small transition rates.

Overall these findings point towards a hierarchical ranking of the three occupation groups. Most importantly, "good" occupations are associated with higher wages and job-stability than the occupational classes ranked below. Furthermore, a typical career does not skip an occupational hierarchy in the sense that there is a jump from unemployment to an occupation other than "Bad" or from "Bad" directly to "Good". However, a large fraction of individuals, having acquired a vocational training degree by construction of the sample, start directly in the mediocre occupation at the time of labor market entry.

The fact that around 50% of downward mobility is associated with wage rises is surprising at first. It points towards mismatch between an individual and an occupation group and is consistent with Borjas and Rosens' (1982) interpretation: Quitting, demotions, promotions and being fired are all a reflection of the same underlying logic, i.e. that more can be earned elsewhere. This is also the logic of the model described in the following where occupational changes are outcomes of optimizing agents.

3. The Model

3.1. Model Structure

The model merges a dynamic discrete choice framework of occupational mobility and endogenous human capital accumulation with a search model and a flexible variance components model of residual wages. It is essentially an extension of Keane and Wolpin's study (1997) of earnings dynamics of young men. Individuals maximize utlility by choosing among four options in each period.[20] The choices - indexed by $j \in \{u, b, m, g\}$ - are unemployment (u) and working in a bad (b), mediocre (m) or good (g) occupation. General and occupation specific human capital accumulation as well as unemployment duration evolve endogenously over time. Individuals, whether employed or not, search for better opportunities and, at a certain rate, draw match effects from all occupations but the one of current employment. Matches are distributed according to some distribution, but they remain constant for the duration of an individual-occupation match. The latter assumption makes identification transparent as the parameters describing the search process are solely identified from the residual variation at the time of mobility. It also implies that a worker who returns to an occupation after an intermediate spell of unemployment or employment in a different occupation draws a new match. Matches are periodically and exogenously broken up, re-allocating the worker into other occupations or into unemployment. Search for better matches induces a jump-process in residuals that is reminiscent of a random walk in occupation-specific skills, with the crucial difference that match quality will be endogenously corrected only in the upward direction. Since matches remain constant during an employment spell and are accepted from other occupations only when they increase earnings, the two types of stochastic processes have very different implications for behavior and welfare.[21] Each individual is "endowed" with a set of unobserved occupation-specific permanent skills, thus allowing workers to differ with respect to comparative advantages. Furthermore, general skills are updated by permanent positive or non-negative shocks in each period, and occupation-specific skills are hit by occupation-specific transitory shocks.

[20] The results reported below are for models in which individuals maximize income in each period. This choice is not necessarily consistent with the solution of a Dynamic Programming problem. The estimation of a full DP-framework, and the investigation of its predictive power, is an extension of this paper. Preliminary results for a number of model specifications with Dynamic Programming are available upon request. Full results will be available soon. As explained below, the results presented should be interpreted as "semi-structural" in the sense that they are derived from a decision-theoretic model in which parameters replicate the transitions of states in the model without solving the complete Dynamic program. For the relationship between the reduced and the structural form in Dynamic Discrete Choice Models, see Hotz and Miller (1993).

[21] Adda et. al (2009) assume that firm-specific matches follow a random walk. Due to the focus of my paper, I allow match effects and the random walk to follow distinct stochastic processes.

This specification merges the discrete-choice framework considered in Keane and Wolpin (1997) with a random-walk model of residual earnings.[22]

I index individuals by i and the time period by t. The optimal choice in period t for individual i is denoted by $j_i^*(t)$. Dummy variables are written as 1(.), equal to one if the condition in brackets is met, and zero otherwise.

3.1.1. Labor Market Frictions and the Process of Job Search

Figure 13 demonstrates the search process a worker is engaged in during unemployment or employment. When entering the labor market, with probability λ_0 an individual gets in contact with the three occupations of potential employment. He then decides if to accept the wage offered or to remain unemployed. With probability $1 - \lambda_0$ he does not have these choices and is forced to stay unemployed. As the career progresses, unemployed workers face the same situation like a labor market entrant. Employed workers continue to receive offers from all other occupations at a rate λ_1, and they are exogenously displaced into unemployment with probability δ. A considerable fraction of occupational switches is associated with wage decreases in the sample. As specified so far, the model cannot generate downward employment mobility. I thus introduce a further source of shocks: At a probability κ an employee cannot stay in his present match, but he can choose among all other options. Given that the worker is not forced into unemployment I refer to this type of shock as a "demotion".[23]

[22]Random Walk models of residual earnings and wages have a long history in labor economics. See for example Hause (1979), MacCurdy(1982), Baker (1997), Storesletten et al (2004) and Guvenen (2009). Some of these studies assume instead that residual earnings follow an AR(1)-process. So far no consensus has been reached as to the correct specification of the residual process. I choose a random walk for computational purposes.

[23]See Jolivet et. al (2006) for an empirical equilibrium job search model with firm mobility introducing this kind of shock for the same reason. They refer to the shock as "re-allocation". Given my focus on vertical occupational mobility instead of firm mobility I chose the notation "demotion"-shock instead.

3.1.2. Occupation Specific Wages

Log-wages in each of the three occupation groups are assumed to be described by linear-quadratic equations that condition on all relevant state-variables determining a worker's skill. Skill prices vary across occupations. This wage structure, which is exogenous by the partial equilibrium nature of the model, allows for great flexibility in the specification of unobserved heterogeneity. The trade-off is that general-equilibrium effects need to be subsumed into skill-prices rather than be explicitely derived from an equilibrium model. Alternatively, one can interpret wages as the outcome of Nash-bargains between employers and employees.

More specifically, an individual's potential wage in occupation j in period t is given by the product of an occupation-specific skill price P_t^j and an occupation-specific skill-index H_{it}^j, with $j \in \{b, m, g\}$. Log-wages are thus given by $p_t^j + h_{it}^j$, where $p_t^j = \ln\left(P_t^j\right)$ and $h_{it}^j = \ln\left(H_{it}^j\right)$. I specify the following parametric log-human-capital functions and log-skill-price functions:

$$h_{it}^j = \alpha_{0,i}^j + \alpha_1^j * x_{it} + \alpha_2^j * (x_{it})^2 + \alpha_3^j * ten_{it}^j + \alpha_4^j * \left(ten_{it}^j\right)^2 + u_{it} + \mu_{it}^j + \varepsilon_{it}^j \quad (3.1)$$

$$p_t^j = \alpha_5^j * t + \alpha_6^j * (t)^2 + \alpha_7^j * U_t \quad (3.2)$$

where t is a linear trend, U_t is the unemployment rate in the current period, x_{it}^j is actual experience, defined as the number of years the individual has spent in the labor force minus the total amount of years spend in unemployment, and ten_{it}^j is occupation-specific tenure. The laws of motion for human capital accumulation are given by the system

$$x_{it} = x_{i,t-1} + 1(j_{i,t-1}^* \in \{b, m, g\}); \quad x_{i0} = 0 \quad (3.3)$$

$$ten_{it}^j = ten_{i,t-1}^j + 1(j_{i,t-1}^* = j); \quad ten_{i0}^j = 0. \quad (3.4)$$

These two equations clarify that both, experience and tenure, are endogenous, starting from a value of zero and evolving consistently with the choices made in each period. No labor market experience is added while being unemployed, as clarified by equation (3.3). Therefore, actual rather than potential experience enters the log-wage equations.

Equation (3.1) specifies occupation-specific skills as second-order polynomials of actual general experience and occupation-specific tenure, both of which are endogenous and evolve with respect to the equations (3.3) and (3.4). Parameters vary freely across occupations. In particular, general experience is allowed to have different returns in the three occupation classes. Standard theoretical models of vertical occupational mobility, such as the Gibbons and Waldman (1999, 2006) model, argue that managerial jobs have higher returns to general experience than jobs on lower ranks. For example, a bank teller might not get much more productive over time, while only individuals with a large amount of labor market experience can manage a group of bank tellers or a branch. The estimates of my model will determine if this assumption is valid for the data and the occupational classifications used herein.

Unobserved heterogeneity is comprised of four components. First, each individual is endowed with a full set of occupation-specific intercepts, $\alpha_{0,i}^{j}$ which are random in the population. These parameters are "innate" in the sense that they determine an individual's comparative advantage and earnings potential at the beginning of a career. Second, occupation-specific skills are hit by transitory shocks ε_{it}^{j}. These two model components together with the specification of the skill indices are very close to the model estimated by Keane and Wolpin (1997). I extent their model by adding two dynamic unobserved skill components, the random walk component u_{it} updating the level of permanent skills in each period, and an occupation-specific match effect μ_{it}^{j}. I further describe the stochastic structure of unobserved heterogeneity below.

Equation (3.2) is a parametric specification of occupation-specific skill prices. Theoretically, they can be identified non-parametrically using year fixed effects. For the sake of interpretation of parameters I choose a parametric skill price function instead, thought to capture general equilibrium effects driven by demand side shocks.[24] In particular, the model allows occupation specific skill prices to trend and to react to aggregate fluctuations. The relative strenght of trends across the occupations reflect structural change, and the unemployment rate allows occupation-specific labor demand to react differently to business cycle fluctuations.[25] Both components introduce exogenous variation into the model and the choice rules of optimizing agents.

3.1.3. Unemployment Benefits

Unemployment is a choice available to a worker at any point in time. The availability of unemployment benefits collected by an individual in the IABS allows me to consider a flexible specification of unemployment insurance which is consistent with the German unemployment system. This system is fairly complicated. It distinguishes between unemployment insurance benefits ("Arbeitslosengeld", AG) and unemployment assistance ("Arbeitslosenhilfe", ALH). AG can be collected only if an individual has worked at least 12 month over the last three years, and only up to a certain amount of time. Afterward, the unemployment benefits drops to the ALH level. The time limit depends on the age of an individual. Rules regarding this time limit have been changed in the mid-eighties, with more severe changes for the elderly who are not present in my sample. I thus assume that over the sample period, the unemployment benefit system has remained constant, and I specify the model in such a way that it is consistent with this assumption. Both, AG and

[24]The model does not solve for general equilibrium. See Keane and Wolpin (1997) for a discussion of a general equilibrium framework with competitive markets and without frictions that can replicate the wage structure assumed here. Alternatively, one can specify a model with competitive Nash bargaining between workers and firms in order to reach at an equation of this form. I feel that this does not add to the discussion and therefore omit it here.

[25]In this paper I abstract from general equilibrium effects due to computational feasibility. Lee and Wolpin (2006) show that the occupational composition of an economy has indeed changed over the last three decades in the US. I have also computed extensive aggregate statistics for the German data, revealing similar trends in Germany. Results are available upon request.

Devereux (2002) shows that occupation classes react differently to business cycle fluctuations in the US. Buettner, Jacobebbinghaus and Ludsteck (2009) replicate the study using the IABS and reach at similar conclusions.

ALH depend on past wages earned on the job.[26]

Let dur_{it} denote an individuals' observed unemployment duration in period t, $w_{i,-1}$ the log-wage observed in the last period an individual was employed - not necessarily $t-1$ - and $1(dur_{it} \geq 2)$ a dummy equal to one if unemployment duration is at least two years long, in which case AG drops to the ALH level. Unemployment benefits are modelled by the following estimation equation:

$$\begin{aligned} b_{it}^{u} &= \alpha_0^u + \alpha_1^u * t + \alpha_2^u * dur_{it} + \alpha_3^u * (dur_{it})^2 + \alpha_4^u * \max\{w_{i,t-1}, 0\} \\ &\quad + \alpha_5^u * dur_{it} * \max\{w_{i,t-1}, 0\} + \alpha_6^u * 1(ALH) + \varepsilon_{it}^u, \end{aligned} \tag{3.5}$$

where unemployment duration evolves endogenously according to the law of motion

$$dur_{it} = dur_{i,t-1} + 1(j_{i,t-1}^{*} = u); \quad dur_{i0} = 0. \tag{3.6}$$

Unemployment benefits are thus determined by a time trend, a quadratic in unemployment duration, the last wage earned and an interaction between unemployment duration and the last wage earned. Unemployment benefits are corrected downward to the ALH level after two years of unemployment. Theoretically, unemployment benefits collected should be a deterministic function of the variables entering equation (3.5). In reality, rules admit several exceptions, and individuals might not collect the actual amount they are eligible to earn. I thus allow for random deviations captured by ε_{it}^u, turning (3.5) into an estimation equation.

Like general experience and occupation specific tenure, unemployment duration evolves endogenously. Equation (3.6) clarifies that it increases only if an individual has remained unemployed in the preceding period.

[26] For a more detailed discussion of the German unemployment insurance system, see for example Fitzenberger et al. (2004), Hunt (1995) and Adda et al (2009a).

3.1.4. Stochastic Structure

The model features several variance components of unobserved heterogeneity. Individuals are endowed with occupation-specific skill levels $\alpha_{0,i}^{j}$ that are hit by transitory shocks ε_{it}^{j} in each period. To significantly reduce the state-space of the Dynamic Programming Problem I follow most applications of Discrete Choice Models and assume that the former discretely distributed. In particular, as clarified by equation (3.7), there are K types of individuals in the population, each endowed with a vector $\left(\alpha_{0,k}^{b}, \alpha_{0,k}^{m}, \alpha_{0,k}^{g}\right)_{k=1,\ldots K}$ of occupation-specific skills. The type-proportions π^{k} sum up to one and will be estimated. The choice of K - the total number of types - is somewhat controversial. A small literature documents problems with conventional likelihood based tests. In the context of duration models, Baker and Melino (2000) show that such tests tend to determine too large a number of types. I follow a parsimonious approach and set $K = 4$ as in Keane and Wolpin. Experimentation with higher numbers does not change the results.

Transitory shocks ε_{it}^{j} allow occupation-specific skills to fluctuate around their means in each period. By definition of specificity I assume that these shocks are uncorrelated across occupations. General skills are updated by permanent shocks, as described by equation (3.9). Initially the alternative-specific intercepts and permanent shocks cannot be separaterly identified. I thus assume that there are no permanent shocks at the time of labor market entry. The dynamics of match heterogeneity is described in equation (3.10), forcing match effects to be constant for the duration of a match. They introduce permanent earnings dispersion among individuals working in the same occupation and otherwise having the same level of occupation-specific skills.

The following set of equations concisely summarizes the stochastic structure:

$$Prob\left(\alpha_{0,i}^{b} = \alpha_{0,k}^{b}, \alpha_{0,i}^{m} = \alpha_{0,k}^{m}, \alpha_{0,i}^{g} = \alpha_{0,k}^{g}\right) = \pi^{k}, \sum_{k}^{K} \pi^{k} = 1, K = 4 \tag{3.7}$$

$$\varepsilon_{it}^{j} \sim N(0, \sigma_{\varepsilon,j}^{2}) \tag{3.8}$$

$$u_{it} = u_{it-1} + \xi_{it}, \left\{ \begin{array}{c} u_{i0} = 0 \\ \xi_{it} \sim N(0, \sigma_{\xi}^{2}) \end{array} \right\}, \tag{3.9}$$

$$\mu_{it}^{j\in\{b,m,g\}} = \left\{ \begin{array}{c} \mu_{i,t-1}^{j} \ if \quad j(t)=j_i^*(t-1) \\ \nu_{it}^{j} \sim N(0,\sigma_\nu^2) \text{ otherwise} \end{array} \right\} \tag{3.10}$$

3.1.5. The Decision Problem

Bellman Equations At the beginning of each period, individuals have to choose among the four alternatives $j \in \{u, b, m, g\}$. They observe their potential alternative-specific log-wages $w_{it}^{j\in\{b,m,g\}} = p_t^j + h_{it}^j$, with p_t^j and h_{it}^j given by equations (3.2) and (3.1) respectively, and unemployment benefits b_{it}^u, given by (3.5). Both, employers and employees are perfectly informed about the current state. Given wages, unemployment benefits and the current individual-specific states, employees choose the alternative with the highest expected pay-off. To concisely summarize the decision problem, let $S_{it} = \left(X_{it}, k, \mu_{it}^j, u_{it}, \varepsilon_{it}^j\right)_{j\in\{u,b,m,g\}}$ be the period-t state vector for individual i, composed of the observable variables X_{it}, the type of the individual k and all components of unobserved heterogeneity. The observable state variables include, aside from the controls entering the wage and unemployment benefit equations, an individual's age. Although age does not directly enter the model, it needs to be included in the state space due to the fact that actual labor market experience and unemployment duration do not add up to potential experience. For example, an individual who is unemployed in the first and the fifth year of his career will have the same actual experience and unemployment duration after five years of experience as someone who works in the first three years, becomes unemployed in year four and is observed at the end ofthe fourth period.

With discount factor β, the choice-specific values are given by

$$\begin{aligned} V_{it}^{j(t)\neq u}(S_{it}) &= \exp\left\{w_{it}^j\right\} + \beta * \lambda_1 * E\left[\max\left\{V_{it+1}^a(S_{it+1}), a \in \{u,b,m,g\}\right\}|S_{it}\right] \\ &+\beta * \kappa * E\left[\max\left\{V_{it+1}^a(S_{it+1}), a \neq j(t)\right\}|S_{it}\right] \\ &+\beta * \delta * E\left[V_{it+1}^u(S_{it+1})|S_{it}\right] \end{aligned}$$

$$
\begin{aligned}
& +\beta * (1-\lambda_1-\kappa-\delta) && (3.11)\\
& *E\left[\max\left\{V_{it+1}^{j(t)}\left(S_{it+1}\right), V_{it+1}^{u}\left(S_{it+1}\right)\right\}|S_{it}\right] &&\\
V_{it}^{u}\left(S_{it}\right) \quad = \quad & \exp\left\{b_{it}^{u}\right\} + \beta * \lambda_0 * E\left[\max\left\{V_{it+1}^{a}\left(S_{it+1}\right), a \in \{u,b,m,g\}\right\}|S_{it}\right] &&\\
& +\beta * (1-\lambda_0) * E\left[V_{it+1}^{u}\left(S_{it+1}\right)|S_{it}\right] && (3.12)\\
S_{it+1} \quad = \quad & \Lambda\left(S_{it}\right) && (3.13)
\end{aligned}
$$

Equation (3.11) is the value function for employment. In the current period, individuals receive wages w_{it}^{j}. In the next period, they draw wage offers from all occupations at a probability λ_1, are forced out of the current occupation but can choose among all the remaining options with probability κ, are forced into unemployment with probability δ, and can choose only among the current occupation and unemployment with probability $(1-\lambda_1-\kappa-\delta)$. Equation (3.12) is the value function for unemployment, an alternative that can be chosen in any period. An unemployed individual collects unemployment benefits b_{it}^{u}, draw offers from all the occupations at a probability λ_0 next period, and do not receive any offers at probability $(1-\lambda_0)$. The final equation is the updating rule of the state space. General experience x_{it}, occupation specific tenure ten_{it}^{j}, unemployment duration dur_{it}, the random walk u_{it} and the match effects μ_{it}^{j} are updated according to the dynamic equations (3.3), (3.4), (3.6), (3.9) and (3.10).

Computational Considerations and Further Assumptions The system of equations forms a standard finite horizon Dynamic Programming Problem and can be solved by backward recursion. A terminal condition for the Bellman equation in the last period of an individuals' life-cycle is required. A common approach is to solve the problem backward from the expected age of retirement, which is usually 63 years in Germany. Due to its dynamic nature, the state space becomes infeasibly large at high experience values. I assume instead that after 23 years of potential labor market experience - the highest value of experience observed in the data - a worker remains immobile for the rest of his career and experiences annual wage growth of 2 percent.

The backward recursion involves the computation of value functions over very large period-specific state-spaces and multi-dimensional numerical integration. To reduce the state space I introduce the following additional restrictions and assumptions:

- I solve the Dynamic Programming Problem only for state space points that are observed in the data or that can be reached by an individual from his current state. This "trick" only works for variables that are observable to an econometrician. Most importantly, the restrictions $\sum_{j=b,m,g} ten_{it}^{j} = x_{it}$ and $x_{it} + dur_{it} = exper\_potent$ decrease the number of state-space dimensions by two. Furthermore, with the oldest cohort entering in 1978 and the youngest cohort entering in 1995, a potential experience of zero is observed only in these years. Likewise, careers with length of twenty-three years are only observed in the data starting in 2000. Since calendar years enter the state-space through the skill-price functions, applying this heuristic significanly reduces the size of the state space.

- Unemployment rates enter the state-space through skill prices. To compute the Bellman equations, individuals need to form expectations about the future. I assume that workers have perfect foresight about next periods' unemployment rate. This avoids to integrate over a stochastic process for unemployment rates. Furthermore, conditional on calendar year, unemployment rates are uniquely determined and thus do not increase the size of the state space at all.

- Transitory skill shocks, permanent skill shocks, and draws from the distribution of match effects, are continuous state variables. While transitory shocks are serially uncorrelated and thus can be easily integrated numerically, the value of a current shock or match provides some information about the distribution of shocks or matches that can be expected in the future. I follow the common approach and discretize the state space over these variables.

Further details regarding the computational procedure can be found in the appendix.

3.1.6. Determinants of Career Progression

In this section I discuss the main features of the model that drive the joint dynamics of earnings and vertical occupational mobility. Similar to earlier studies, most notably Keane and Wolpin (1997), human capital accumulation directly influences mobility decisions. General human capital, having differential returns in different occupations, can induce individuals to switch after some time if, on average, occupations with higher returns are associated with lower intercepts. This is, in fact, the cornerstone of Gibbons and Waldman's (1999, 2006) theory of vertical occupational mobility. In their model, individuals can earn relatively much early in a career in low-ranked occupations, but additional experience is of little value. Managerial and professional occupations in contrast, with low payoffs early in the career, have large returns to general human capital. Since I do not impose any ex-ante restrictions on parameters, my estimates below will determine if these assumptions are met in the data used herein. Occupation-specific human capital, which becomes useless when switching occupations, represents a fixed cost to mobility. Other than human capital endowment, transitory shocks to occupation-specific skills are the only component of the Keane and Wolpin model specification generating mobility. Since these shocks are assumed to be iid-random variables, they generate quite erratic mobility patterns early in the career when occupation-specific human capital does not bind as a fixed cost to mobility yet. In contrast, once individuals have accumulated enough specific capital later in their career, mobility rates become very low. Furthermore, since transitory shocks are the only source of residual wage fluctuations, they serve the dual role of explaining any unexplained variation in both, mobility and earnings. Consequently, the model is to restrictive for the analysis of the the joint long-run evolution of earnings and mobility. To fix these shortcomings, I add two variance components.

First, individuals differ with respect to the quality of their match, but they can search for better opportunities in the other occupations. Like transitory shocks, this process has a dual role. On the one hand-side, due to the incentive to look for better alternatives, it induces mobility. Individuals with a relatively bad match are most likely to be mobile, a prediction I will use below to test if the estimated extend of match heterogeneity is biased by unobserved within-type general skill heterogeneity. On the other hand, and more importantly, match effects generate inequality among individuals who belong to the same type, have the same human capital stock and are employed in the same occupation. This inequality is permanent for the duration of a match and endogenously determines the quality of matches accepted by individuals. Match heterogeneity and the possibility to search for better matches induces a jump process in residual wages. Jumps are observed whenever an individual changes occupations. As noted in the introduction, when not controlling for the systematic relationship between mobility and residual earnings fluctuations, this earnings variation will be, at least to some extent, included in the estimates of type heterogeneity.

Second, individuals are hit by permanent shocks to their general skill level. Consequently, over time individuals of the same type progressively differ with respect to their average earnings potential. The literature on earnings dynamics stresses the important role of this variance component for the welfare evaluation of earnings fluctuations. Since permanent shocks are uninsurable, social insurance can improve welfare. Incorporating permanent shocks into a structural model makes at least three contributions. First, since earnings intercepts can be interpreted as initial conditions to a unit-roots process, not accounting for permanent updates can significantly alter estimates of the importance of innate abilities. Second, permanent shocks, though neutral with respect to occupational mobility, can influence the choice of an individual to be employed or not. In particular, individuals who are hit by a negative permanent skill-shock can be driven into unemployment, and individuals who are hit by a very good shock are less likely to become unemployed. Not accounting for this selection mechanism can potentially bias the role of permanent shocks downward. Only very few studies, most notably Altonji et al. (2009) and Low et. al (2009), correct for this bias, albeit in a reduced form framework. Third, the model allows for the estimation of a unit roots shock "cleaned" from residual and permanent wage variation that is intrinsically linked to occupational mobility. Indeed, as clarified above, search for matches induces a jump-process in residual earnings at the time of occupational mobility. When estimating unit-roots models without controlling for mobility, residual wage fluctuations that induce behavioral adjustments are mistakenly interpreted as exogenous and neutral shocks, therefore leading to an overbias of random walk shocks.

3.2. Further Discussion of Model Features

The literature on earnings and mobility dynamics is large and still growing. The majority of work considers either simple specifications of mobility and a flexible specification of residual wage dynamics, or vice versa. It is the central contribution of this paper to formulate a model of the joint dynamics of earnings and mobility. The rich specification of unobserved heterogeneity comes at the cost of keeping the model in a partial equilibrium framework. Equilibrium search models, in which the distribution of match effects is endogenous, are an interesting alternative. However, to be computationally tractable, these models can admit neither serially correlated errors, nor non-stationarity. Both dimensions are important to test the quantitative importance of different sources of earnings fluctuations and innate ability.

Unlike other discrete choice models with search, most notably Low et al. (2009), I assume that workers draw new matches only at some constant rate rather than a flexible function of observed variables such as age or tenure. Conceptually it is relatively straightforward to extend the model in this direction. I do not choose to do so for two reasons: First, introducing a function $\lambda(.)$ in terms of experience and other observable variables improves the fit of the model to life-cycle mobility patterns by construction without adding economic structure. Second, both the function $\lambda(.)$ and the parameters governing the dynamics of the match effect "target" life-cycle mobility patterns, potentially rendering identification and interpretation difficult. In particular, specifying a flexible function $\lambda(.)$ introduces duration-dependence in employment states by assumption. Rather, in my model specification duration dependence is an endogenous outcome of optimizing agents.

I also do not follow the convention of allowing for non-pecuniary alternative-specific benefits and fixed costs to mobility. This class of parameters can siginificantly improve the fit of Dynamic Discrete Choice models, but at the cost of pushing the model out of a pure human capital framework. The assumption that there are no non-pecuniary benefits for unemployment spells is more controversial. It implies that search for a new job is equivalent to full-time employment.

4. Estimation

Most applications of discrete choice modelling in Labor Economics model only choice probabilities, usually using linear probability models or Probit, thus restricting the number of parameters that can be identified and making exclusion restrictions necessary. In contrast I derive the joint simulated likelihood of the whole life-cycle profile of wages, optimal choices and observables, $\{w_{i,t}, X_{it}, j_i^*(t)\}_{i,t}$, where $X_{it} = \left(t, U_{it}, x_{it}, ten_{it}^{j\in\{b,m,g\}}, dur_{it}, w_{i,-1}, age_{it}\right)$ is the vector of predetermined variables. To spare on notation, I also define $\mu_{it} = \left(\mu_{it}^j\right)_{j\in\{b,m,g\}}$ to be the full vector of match effects, $\widetilde{S}_{it} = \left(X_{it}^j, k, \mu_{it}^j, u_{it}\right)_{j\in\{u,b,m,g\}}$ to be the state-vector excluding transitory shocks to skills, and $\widehat{w}_{it}^j$ to be the wage in alternative j predicted by the conditioning variables. For ease of exposition I briefly discuss the likelihood in a model without frictions in the next section. Likelihoods for the general model can be found in the appendix.

The estimation is involved and cannot be carried out using standard estimation packages. To investigate the numerical properties of my estimation routine, to explore the strength of parameter identification, and to rule out coding mistakes, I have tested each program using extensive Monte-Carlo analyses before applying it to actual data. Programs performing the Monte-Carlo analyses are available on request.

4.1. The Likelihood

The computation of the likelihood function starts with the decomposition of the individual i, period t likelihood contribution conditional on $\widetilde{S}_{it}$,

$$L_{it} \equiv \Pr(w_{it}^{j^*}, j_i^*(t) \mid \widetilde{S}_{it})$$

$$
\begin{aligned}
&= \Pr(w_{it}^{j^*(t)} \mid \widetilde{S}_{it}) * \Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}) \\
&= \Pr(j_i^*(t) \mid \widetilde{S}_{it}) * \Pr(w_{it}^{j^*(t)} \mid j_i^*(t), \widetilde{S}_{it}). \qquad (4.1)
\end{aligned}
$$

Given that $\Pr(w_{it}^{j^*(t)} \mid j_i^*(t), X_{it})$ is the conditional density of a continuous random variable, while $\Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, X_{it})$ is a discrete object, I will use the decomposition in the second line of (4.1) to avoid simulation of an object with measure zero. The only source of randomness in this conditional likelihood are the transitory occupation-specific shocks to skills. By assumption, they are uncorrelated across alternatives. Bellman-equations enter the likelihood through conditional choice probabilities. Since they are solved in terms of wages, rather than log-wages, computing choice probabilities involves exponentiation. Denoting the expected payoff next period when choosing j today - given by the expressions multiplied by β in equations (3.11) and (3.12) - as $V_{it+1}\left(\widetilde{S}_{it}, j\right)$ yields $V_{it}^{j^*(t)}\left(w_{it}^{j^*(t)}, \widetilde{S}_{it}\right) = \exp\left(w_{it}^{j^*(t)}\right) + \beta * V_{it+1}\left(\widetilde{S}_{it}, j^*(t)\right)$ and $V_{it}^k = \exp\left(\widehat{w}_{it}^k\right) * \exp\left(\varepsilon_{it}^k\right) + \beta * V_{it+1}\left(\widetilde{S}_{it}, k\right)$. Taking use of the fact that the exponentials of independent Normally distributed random variables are independent we reach at:

$$
\begin{aligned}
\Pr(w_{it}^{j^*}, j_i^*(t) \mid X_{it}, \mu_{it}, \alpha_i, u_{it}) &= \phi\left(\frac{w_{it}^{j^*(t)} - \widehat{w}_{it}^{j^*(t)}}{\sqrt{var\left(w_{it}^{j^*(t)} - \widehat{w}_{it}^{j^*(t)}\right)}}\right) \qquad (4.2) \\
&* \prod_{k \neq j^*(t)} \Phi\left[\left(\frac{1}{\sigma_{\varepsilon,k}}\right) * \left(\ln\left\{ \begin{array}{c} \exp\left(w_{it}^{j^*(t)}\right) + \\ \beta * \left[V_{it+1}\left(\widetilde{S}_{it}, j^*(t)\right) - V_{it+1}\left(\widetilde{S}_{it}, k\right)\right] \\ -\widehat{w}_{it}^k \end{array} \right\} \right)\right]
\end{aligned}
$$

where ϕ is the pdf and Φ is the cdf of the Standard Normal Distribution. The first term is the conditional wage density and the product term is the conditional choice probability. In the case with $\beta = 0$, this probability reduces to standard Probit choice probabilities with uncorrelated errors.

The computational difficulty arises because the remaining part of unobserved heterogeneity, μ_{it}, α_i and u_{it}, needs to be integrated.[27] In particular, given the discreteness of α_i we have

$$L_i = \sum_k \pi_k \left\{ \prod_t \left[\int L_{it}(X_{it}, \mu_{it}, \alpha_i, u_{it}) f(\mu_{it}, u_{it} | \mu_{it-1}, u_{it-1}, X_{it}, \alpha_i) d\mu_{it-1} du_{it-1} \right] \right\}. \quad (4.3)$$

Simulation of the term in curly brackets is unavoidable because $f(\mu_{it}, u_{it}|\mu_{it-1}, u_{it-1}, X_{it}, \alpha_i)$ cannot be factored into $f(u_{it}|u_{it-1}) * g(\mu_{it}|\mu_{it-1})$. In other words, even conditional on the past state of the random walk and the match effect, u_{it} and μ_{it} are not independent from the other pre-determined variables. To understand why, compare two individuals with the same past state, but a different current realization of the random walk. The model predicts that the individual with the higher value is more likely to stay employed and thus to accumulate more human capital. Since general experience is part of the observables this generates a relationship between u_{it} and X_{it}. A similar argument applies to the endogeneity of μ_{it-1}. Consequently, the likelihood of an individual's whole labor market history needs to be simulated. Since the draws from the distributions are continuous while the associated variables in the state-space of the Bellman Equations are discretized, I interpolate the value functions between state-space points. Details can be found in the appendix.

Denoting Θ to be the parameter vector, the estimates are given by the maximand of the log-likelihood:

$$\widehat{\Theta} = \arg\max \sum_{i=1}^{N} \log L_i(\Theta). \quad (4.4)$$

Standard errors are computed from the inverse Hessian.

[27] It is important to note that given an individuals decision the non-stochastic nature of the laws of motions for experience, tenure and unemployment duration implies that their transition probabilities are equal to one. This is different from many applications in Industrial Organisation, where the observable state variables evolve stochastically, even conditional on choices. See e.g. Rust (1987).

4.2. Some Numerical Issues

I use a nested algorithm that simulates the Dynamic Programming Problem and the likelihood in the inner step given the current value of the parameter space, and that updates the parameter space in the outer loop following a standard maximization routine.[28] Due to the discreteness of the choices, many simulators of choice probabilities are discontinous as well, introducing numerical difficulties when maximizing the simulated likelihood. Numerical integration in Discrete Choice Models can thus become potentially involved, even when the specification of unobserved heterogeneity is much simpler than considered in my model.[29] In Probit models the standard approach to "smooth" the likelihood is given by the Geweke-Hajivassiliou-Keane-simulator. Given my assumption that transitory occupation-specific shocks are uncorrelated across alternatives, this simulator reduces to a simple product of cdf's, as displayed in equation (4.2). Consequently, this assumption, beside being economically plausible by the definition of "specifeity", has the advantage that transitory shocks can be integrated in closed form, and that the remaining numerical integration (4.3) takes place over a continous integrant $L_{it}(X_{it}, \mu_{it}, \alpha_i, u_{it})$. Hence, the simulated likelihoods are continous as well, enabling me to use fast gradient-based maximization methods of the simulated likelihood.[30]

Equation (4.2) demonstrates that the Bellman-equation enters the likelihood as a variable that depends on the vector of predetermined variables. I thus compute the Bellman equation and match it to the vector of state-variables observed for an individual. The current simulated state of the random walk process and the draws of the match effects are treated as if observable and matched to the respective Bellman equation. Since draws are continous, while its state-space is discretized, I extrapolate linearly between the state

[28]I use a Broyden-Fletcher-Goldfarb-Shanno (BFGS) maximization algorithm.

[29]For an excellent discussion of numerical methods in Discrete Choice Models, refer to Train (2003). For a more general treatment of simulation-based estimation, see Gourieroux and Monfort (1996), and for a discussion of practical issues see Mariano, Schuermann and Weeks (2000). General numerical methods are described in Judd (1998).

[30]The model has a number of linear constraints on probability parameters to be estimated. One can easily use transformations of parameters such that these conditions are automatically met. This avoids using potentially much slower non-gradient based methods.

points.

Monte-Carlo studies and actual estimation I have conducted assured that the likelihood has numerically desirable properties. Specifically, it converges to the same maximum from any economically plausible starting point of parameter values, and it is very insensitive to the number of draws taking in the simulation once a certain threshold has been passed.

4.3. Identification

The model is kept tightly parameterized, and identification is transparent. In total there are 54 parameters, 29 parameters of which describe the observable part of log-income and are essentially selection-corrected regression estimates. Only 2 parameters capture the processes of the random walk and the match heterogeneity of the unobservable part, while 15 parameters determine the non-parametric distribution of comparative advantage, and 4 parameters determine the variances of the transitory shocks. The remaining four parameters describe the search process.

To understand how the regression parameters are identified it is instructive to look at equation (4.2). The first term is the likelihood of a regression model that estimates the parameters in equations (3.2), (3.1) and (3.5) separately and without selection correction. The term in brackets recognizes that the choice of occupation and employment status is endogenous and corrects for potential selection biases. This term is fully consistent with the model's theoretical structure. For example, a large literature tries to identify returns to occupational tenure, but recognizes that these estimates are plagued by selection biases. In particular, estimates might be overbiased because it just happens that over time only individuals who have found a particularly good match in this occupation remain there, or because they have comparative advantage in this occupation. The model is fully consistent with these types of selection biases. It chooses the regression parameters in such a way that net of match effects and occupational skill shocks it remains optimal to be in this occupation. Two exogenous sources of variation, occupation specific time trends and differences in the sensitivities to business fluctuations, help further to identify the parameters.

Identification of the parameters describing the processes of match effects is particularly transparent. For an individual, the wage change between two periods is given by

$$
\begin{aligned}
w_{it}^{j^*(t)} - w_{it-1}^{j^*(t-1)} &= X_{it}^{j^*(t)}\beta^{j^*(t)} - X_{it-1}^{j^*(t-1)}\beta^{j^*(t-1)} \\
&\quad +\xi_{it} + 1\left[j_i^*(t) \neq j_i^*(t-1)\right] * \nu_{it}^{j} + \varepsilon_{it}^{j^*(t)} - \varepsilon_{it-1}^{j^*(t-1)} \qquad (4.5)
\end{aligned}
$$

so that a change in the match effect is only observed at the time of mobility. Therefore, it is the systematic residual wage variation at the time of mobility that identifies the parameter determining the distributions of match effects. Since mobility takes place only when there is an improvement in match quality, this distribution is truncated, a fact which is taken care off by the conditional choice probabilities. Distributional assumptions are required to infer the whole distribution from the truncation.

The relative importance of permanent versus transitory shocks is identified from the remaining residual wage fluctuations. A unit roots process predicts that these variances increase linearly in experience. Consequently, the variance of the permanent shocks is chosen to match a linear trend of variances to the empirical residual variances.

Heckman and Singer (1984) are the first to introduce non-parametric estimation of type proportions in the context of duration models. It is an attractive choice for models that lend themselves to an interpretation in terms of a finite number of groups. It is also an attractive choice in any model that requires multi-dimensional numerical integration or a Dynamic Programming Problem that needs to be solved for each type. In my model there are four types, each of which is associated with a vector of occupation-specific intercepts $\left(\alpha_{0,k}^{b}, \alpha_{0,k}^{m}, \alpha_{0,k}^{g}\right)$ and a discrete probability π^k. Both sets of parameters are estimated. To gain some intuition for the identification of these parameters it is important to note that it is theoretically possible though computationally intractable to estimate the model for each individual. The Panel nature of the data together with the choice-probabilities allow to estimate occupation-specific intercepts for each individual in the sample. Consequently, one could then plot the distribution of these fixed effects and match a non-parametric function. This function is thus non-parametrically identified. Allowing for k types of fixed effects instead and estimating the associated probability masses is an extreme way of discretizing this distribution.

Search frictions essentially enter as scale parameters on transition rates between occupations and between employment and unemployment. Consider for example the transition rates from employment into unemployment, and suppose that for any variable determining wages the estimated wage equations predict that it is optimal to remain employed. In this case, the parameter δ - the probability that a match is broken up - "matches" the observed transition rates from employment to unemployment. Similarly, the probability of a demotion-shock κ is chosen to match the transition rates between occupations which are accompanied by wage decreases. In that sense, the probability that a worker is hit by a displacement or a demotion shock is identified from transitions which should not take place given the estimated predicted wages. In contrast, job search parameters λ_0 and λ_1 scale down upward transitions the model predicts to take place given predicted wages in the different choices but are in fact not observed. For example, suppose that given and individual's observables the model predicts he should be highly likely to switch occupations, but is observed not to do so. In this case λ_1 scales down the conditional choice probability to be consistent with mobility not taking place. The same intuition holds for transitions between unemployment and employment which are predicted to be optimal, but that are not observed in the data.

5. Results from Dynamic Probit Models

Dynamic Probit Models, implicitly assuming that individuals are income rather than wealth maximizing, are a natural starting point for the estimation of Dynamic Discrete Choice Models. Although they are based on an explicit individual optimization problem the resulting decision rules are reduced forms of dynamic policy functions.[31] Consequently, results should be interpreted as "semi-structural". Given the focus on earnings and mobility dynamics and given preliminary results, conclusions of the paper are not expected to be significantly altered when incorporating the full Dynamic Programming solution into the estimation.[32] However, numerical policy experiments that potentially change individuals' expectations about the future should not be conducted.

5.1. Parameter Estimates

Parameter estimates of the model are shown in Appendix Table 1. Panel A of the table displays the estimates of the parameters entering the observed part of the log-wage equations. Although occupation-specific intercepts differ across types they are shown in the same Panel as well. Panel B provides parameter estimates describing unobserved heterogeneity and search frictions.

With the exception of some coefficients in the unemployment benefit equation, parameters are precisely estimated and highly significant.[33] The inflation-adjustment of social insurance payments is reflected in the insignificant linear trend estimate of unemployment benefits. Since ALG is not duration adjusted until it drops to the ALH level, unemployment duration does not have a significant impact either. Surprisingly, the discrete adjustment after two years of unemployment is estimated to be insignificant, possibly because very

[31] A large and still growing literature in Industrial Organization, initialized by Hotz and Miller (1993) and motivated by the need to reduce the computational burden of Dynamic Discrete Choice Models, investigate the relationship between Dynamic Programming Models and their reduced form.

[32] Given the complexity of the model and the size of the data, solving the DP program is computationally demanding and has large computer running times even on 64-bit multicore processors. Results for the Keane-Wolpin specification with Dynamic Programming are finalized and available upon request. The conclusions derived from this specification are unaltered when incorporating the Dynamic Programming framework.The estimation of the full model with Dynamic Programming is still in progress. Results will be available soon.

[33] They are also robust to model specifications. I have estimated a large number of models, all of which are a nested versions of the full model considered here. I have started with the simplest model - a Roy-model without type-heterogeneity - and then added pregressively more model features. Most parameters on observables are statistically unchanged across specifications. An exception are the returns to general human capital that adjust to match selection behavior through conditional choice probabilities. I have not listed the parameter values to keep the table transparent. They are available upon request.

few individuals in the sample remain unemployed for such a long time, and because the non-linear terms in duration fit the empirical profile sufficiently well.

There is a strong negative association between occupation-specific slopes and returns to general experience. The bad occupation has the highest intercepts, but the smallest returns to general experience, while the opposite is true for the good occupation. Returns to experience are 2.8 percent in the bad and mediocre occupation, and 6.8 percent in the good occupation. Returns to tenure are very low and below 1 percent. Estimates are quite close to those from simple OLS-regressions (not shown in the table) in which the overall return to experience is 2.8 percent and the return of tenure is 0.6 percent. The estimated tradeoff between intercepts and slopes across occupations satisfies the assumptions used in Gibbons' and Waldman's (1999, 2006) model of promotion dynamics. Consequently, mobility dynamics will follow their predicted theoretical pattern: On average, individuals start in bad or mediocre occupations but eventually move to the good occupation.

The specification of the skill-price functions, capturing economy-wide labor demand patterns, allow for occupation-specific non-linear trends and occupation-specific sensitivities to business cycle fluctuations. Consistent with the Skill Biased Technological Change Hypothesis, good occupations experience the strongest, and bad occupations experience the weakest trends. The sensitivity to business cycle fluctuations is different across occupations as well. Good occupations are the least, and bad occupations are the most sensitive to business cycle fluctuations. This suggests potentially large long-run career effects of business fluctuations, a hypothesis to be tested in future extensions of this work.

Market frictions are large. Although the model estimates that 78 percent of unemployed workers contact an employer, only 3.3 percent (the sum of estimates for λ_1 and κ) of employed individuals draw offers from different occupations. The job breakup rate of 2.5 percent is in accordance with the low estimates reported in Adda et. al (2009) for the same education group in the same data.

Both, a large difference of contact rates between employed and unemployed workers and a relative high contact rate for unemployed workers have been reported in the equilibrium search literature. Compared to the estimates from models of firm-mobility - as documented in work by Postel-Vinay and Robin (2002) and Adda et. al (2009), among many others - the contact rates to employed workers are very low. Given that the model does an excellent job in matching transition rates, firm mobility within occupational classes must be quite high. Consequently I might miss a large part of earnings fluctuations due to a dimension of mobility that is not explicitely incorporated into the model. However, keeping in mind that the variances of both, unit roots shocks and transitory shocks, are identified off the variation within occupational matches, their relatively small estimated values speak against wage changes from firm mobility within occupational classes to be large. The dispersion of match heterogeneity is five times as high as the standard deviation of unit roots shocks, and more than twice as high than the standard deviation of purely transitory shocks. Further evidence against the importance of firm-mobility within occupational matches is given by the large estimated variation of match heterogeneity itself. Search models of firm mobility with high contact rates predict individuals in bad firms to catch up very quickly to their peers in better matches. Since matches in my model are averages over all firms an individual is employed in during an occupational match, search models predict very small dispersion.[34]

5.2. Model Match

Appendix Tables 2 and 3, and Appendix Figures 1 to 11 investigate the match of the model to the stylized facts listed in section 2 and compare it to the fit of simpler specifications. The nested specification refered to as the Keane-Wolpin specification contains type heterogeneity and transitory shocks to occupation-specific skills as the only components of unobserved heterogeneity. The full model without search frictions - a Keane-Wolpin model enhanced with match heterogeneity and a random walk - is considered as well. All models exclude fixed costs to mobility and non-monetary utility components, both of which could further improve the fit of the model.

[34]Hornstein, Krusell and Violante (2006) point out the inability of equilibrium search models to generate large residual wage dispersion. Most of the inequality in empirical Burdett-Mortensen models is driven by the estimates of productivity dispersion.

The full model matches the wage-experience profiles exceptionally well, but somewhat overpredicts the growth of variances. The life-cycle profile of the employment share of the good occupation, a stock variable, is matched almost perfectly. Unemployment rates are initially much too high, an outcome driven by the assumption that labor market entrants face the same situation like unemployed workers. They fall to the actual level after two years and closely follow the empirical profiles afterward. The model overpredicts the share in the bad occupations, but underpredicts the share in mediocre occupations. Wages observed in both alternatives are too similar to generate the large differences in employment shares observed in the data. Differences in non-pecuniary benefits are a candidate to improve the fit in this dimension. Given that levels rather than the evolution over the life-cycle are mismatched, it should not affect the results of this paper.

Appendix Figures 5 to 8 document an excellent fit of the model in terms of transition rates. Given that identification of the distribution of match effects comes from transitions, this is very important. In general, the model replicates the re-allocation of labor from the two lower occupational classes into the good occupation. The only transitions that are not explained very well are transitions between unemployment and good occupations and vice versa. Large wage gains in the good occupation to general human capital attract individuals with much experience. In the opposite direction, since job breakup rates do not vary across occupations, the model predicts to many transitions between good occupations and unemployment.

To appreciate the match of the model, it should be compared to the outcomes of the two simpler specifications. The Keane–Wolpin model, esentially a dynamic Roy model, matches wage levels and variance profiles equally well, although the former has a counterfactual S-shape rather than being concave. However, transition rates between occupations are grossly overpredicted, and the employment share in good occupations approaches 100 percent. Similar problems have been documented by Keane and Wolpin. The underlying cause is the dual role of transitory shocks to serve as wage residuals and to match mobility not predicted by the observables. As a consequence, if wage residuals are large, the model generates high mobility rates. This problem is commonly addressed by introducing non-monetary fixed costs to switching occupations. I choose a different approach by introducing two additional variance components and search frictions into the Roy-model. A model with these components achieves a very good fit in most dimensions of the data. As shown by the results from the full model without search frictions, introducing match heterogeneity significantly improves the fit by breaking up the direct link between residual wage fluctuations and mobility that is present in the Roy-model.

The full model without search frictions also underperforms in many dimensions. Most importantly, since current matches are never broken up and offers from other occupations are drawn in every period, the model grossly overpredicts wage growth and the rise of inequality over the life-cycle. Consequently, unemployment rates quickly reach zero percent. Hence, there must be market frictions that slow down the search process.

Appendix Table 2 documents the empirical and predicted relationship between mobility and wage changes. The full model with frictions matches the empirical fractions of occupational upgrades that are associated with wage increases, and the fraction of occupational upgrades that are associated with wage decreases, quite well. Surprisingly, it underpredicts the fraction of occupational downgrades that are related to wage increases and overpredicts the fraction of occupational downgrades that are related to wage decreases. An exception are occupational changes between mediocre and bad occupations, a result explained by the similar wage structures observed in these two occupations. The Roy-model is quite symmetric with respect to wage changes and occupational changes. For example, the fraction of occupational upgrades and downgrades associated with wage increases and decreases are quite similar, an outcome of mobility pre-dominantly being driven by iid-transitory shocks to occupational skills which generate erratic mobility patterns early in the career.

Appendix Figure 10 plots kernel estimates of the empirical distribution of total life-cycle earnings, refered to as wealth. The predicted wealth distributions are replicated in Figure 11.[35] Given the decrease in sample size at high levels of experience I define a life-cycle to be 15 years long. Life-cycle earnings are cleaned from cohort effects. The full

[35] I would have prefered to plot the kernel densities in the same graph. The density has to be estimated from the individual-level data rather than some aggregate. Estimation thus has to be performed on computers at the IAB. In contrast, all simulations have been performed outside of the IAB.

model replicates the quantitative and qualitative features rather well. In particular, the distribution is slightly skewed with a rather large mass at the upper end of the distribution. The predicted distribution puts slightly too much mass on the lower end of wealth. The prediction from a model without frictions but with match heterogeneity and unit roots shocks is quite similar, with a fatter tail at the upper end of the outcome variable. The simple Roy-model generates a wealth distribution with four maxima, exactly one maximum per type. Within each type, the distribution is rather slim. Consequently, this version of the model fails to generate a good fit to the wealth distribution.

Appendix Table 3 documents the empirical match to transition matrices of earnings. Consistent with the data the full model features a low degree of earnings mobility in two sub-sequent years. The level is somewhat underestimated. Five- and ten year transition matrices are matched remarkably well. In contrast, the Roy-model grossly overpredicts mobility. Again, this is an outcome of transitory shocks representing the only source of residual earnings dynamics.

To summarize, I have investigated the match of the model to many data moments and have found that it is, in most dimensions and in contrast to the simple Roy model such as considered in Keane and Wolpin, remarkably well. Introducing just six additional parameters, one for the unit roots process, one for match heterogeneity and four for labor market frictions, improves the fit of the Roy model significantly and works toward a realistic representation of the dynamics of earnings and vertical occupational mobility.

5.3. The Role of Types

Before turning to the analysis of inequality it is helpful to characterize the systematic differences in career progression for each of the four types in the model. Figure 14 plots type-specific earnings-experience profiles. Types 1 and 2, associated with the blue and red line respectively, strictly dominate the profiles of types 3 and 4.[36] Consequently,

[36]As noted in Cameron and Trivedi (2005, p. 625), "all finite mixture models are unidentified in the sense that the distribution of the the is unchanged if the subpopulation labels are permuted". Applied to

the former group also dominates the latter in terms of wealth, as depicted in Figure 17. Compared to the corresponding densities computed from the Keane-Wolpin model and shown in Appendix Figure 11, the distributions of the low- and high-earnings types have large overlaps, with low- and high-earners observed in each type-group. This demonstrates the Roy-models' propensity to label any individuals with a particularly good match as high-types although their success is not necessarily related to their innate comparative and absolute advantages.

Turning to the analysis of career progression in terms of occupational choice in Figures 15 and 16 surprisingly reveals that type1-employees are predominantly employed in bad and mediocre occupations. With their comparative advantages in these occupations estimated to be very large, they re-allocate toward the good occupations only when drawing very good matches. Type3-employees are concentrated in these two occupations as well, but due to differences in absolute advantages, they receive much lower than type 1. Although experiencing a stronger re-allocation into the good occupation, they remain at the bottom of the wage distribution. Types 2 and 4 experience the highest average earnings growth over their career, the first group because they have a comparative advantage in the good occupation which has large returns to experience, and the second group because they quickly re-allocate from the bad and mediocre occupation into the good occupation.

These results demonstrate that individuals with a similar set of comparative advantages can differ substantially and permanently with regard to their earnings levels. Hence, policies targeting general skills can foster earnings potentials and change career progression even when keeping comparative advantages constant. More concretely, high-school education can generate similar results than government sponsored and occupation-specific apprenticeship programs in terms of long-run career outcomes.

---

my paper, types are not clearly ranked, and one can change the names of the types without changing the models' estimates and predictions.

6. Sources of Life-Cycle Income Inequality

In this section I present the results from numerous counterfactual exercises that investigate the numerical impact of different model components on life-cycle earnings inequality and earnings mobility. I use life-cycle earnings, in the following refered to as "wealth", as the primary outcome of interest since they are commonly thought to better capture differences in welfare than period-by-period earnings. Earnings mobility as defined by the probability that an individual in the p-th quantile of experience-specific earnings distributions is observed in the q-th quintile some time later is another outcome policy makers are interested in. Poverty traps, a situation in which an individual is permanently poor, is a career outcome of particular public interested. In contrast to single-equation models of earnings dynamics, my model features many sources of state-dependence in regards to shocks individuals receive, how skills are valued, and how workers adjust their behavior to different sources of shocks. It is thus well suited for the study of income mobility.

All counterfactual experiments are conducted by simulating a complete life-cyle career trajectory for 10,000 individuals using the actual parameter estimates, but with one set of parameters adjusted to reflect the counterfactual exercise in question. For example, to quantify the role of labor market frictions I eliminate them by setting $(\lambda_0, \lambda_1, \kappa, \delta) = (1, 1, 0, 0)$. All simulated data replicate the cohort structure of the true data used in the estimation routine.

6.1. Sources of Wealth Inequality

I start with computing the percentage changes in standard deviations of the wealth distributions when perform counterfactual experiments. Results are displayed in Table 4. Each row refers to a different counterfactual experiment, and each column refers to a different model specification. I consider the Roy-model with type heterogeneity as estimated in Keane and Wolpin (1997), the full model with search frictions, and a single-equation regression model of log earnings on the explanatory variables of the model, with residuals assumed to be composed of a random walk and purely transitory shocks. Unlike the Keane-Wolpin model, the regression model is not nested within the behavioral model. However, it is the most prominent among single-equation models of residual earnings dynamics and has gained a lot of attention in the literature calibrating heterogenous Dynamic General Equilibrium models to aggregate wealth distributions. I therefore present its estimates for the purpose of comparison. To rule out that differences of results from counterfactual exercises are driven by the estimation method, I estimate the single-equation model by Maximum Simulated Likelihood.[37]

As shown in row one of column one, eliminating type heterogeneity from the Keane-Wolpin specification would reduce the variance of the wealth distribution by almost 74 percent. Another measure of the importance of type heterogeneity for life-cycle inequality is the fraction of wealth inequality that is explained by type fixed effects. This is the measure used in Keane and Wolpin, but it cannot be used for model components that vary over time, such as match heterogeneity and transitory or permanent shocks. Results from such variance decompositions are listed in the lower panel of Table 4. In the Roy-model, 91 percent of life-cycle variation is explained by type heterogeneity, a striking result that already has been anticipated by Appendix Figure 11 with its multi-peaked density. The quantitative impact of type heterogeneity estimated from the German data is almost identical to Keane and Wolpin's estimates from the NLSY. They document that 90 percent of earnings variation is explained by type heterogeneity. Strictly speaking, estimates from simple dynamic Roy-models imply that over 90 percent of wealth inequality is determined even before individuals enter the labor market.

Column two lists the results from counterfactual experiments when using the full model with frictions. The conclusions are strikingly different: When eliminating type heterogeneity, wealth inequality decreases by 34 percent only, almost a half of the number from the Keane–Wolpin model. Similarly, only 41 percent of the earnings variation is driven by type heterogeneity. Furthermore, excluding transitory shocks from the full model does not have

[37]It is common, and much faster, to estimate single equation models by GMM. See for example my discussion in Hoffmann (2009).

an impact on inequality at all, while the corresponding impact in the Keane-Wolpin model is a reduction of inequality by 14 percent. Ruling out match heterogeneity, an element of unobserved heterogeneity that is absent from the Keane-Wolpin specification, would reduce inequality by 27 percent. Hence, the quantitative implications of type and match heterogeneity are estimated to be very similar.

I also plot the impact of counterfactual experiments on the full distribution of wealth in the full model. Results are provided in figures 18 and 19, with the blue line corresponding to the wealth distribution before conducting the counterfactuals. As hinted at by Table 4, the exclusion of transitory and permanent shocks have very small effects on the full distribution. In contrast, exclusion of type heterogeneity alters the shape of the distribution by reducing its mean, dispersion and skewness. Although I use the average for occupation-specific intercepts over types to avoid level effects, the average wage still decreases, an effect driven by the elimination of comparative advantages. The distribution, now close to being symmetric, still exhibits a fat tail at the upper end. This probability mass represents individuals who find good matches in the good occupation before the fixed costs of mobility derived from occupation-specific human capital accumulated in other occupations with lower returns ot experience become high enough to prevent individuals to switch.

Figure 19 demonstrates the interesting effect that excluding match heterogeneity exacerbates the role of type heterogeneity. Without search for better matches, the distribution becomes bi-modal, with one group "stuck" in low earning jobs, and one group being associated with high earnings, a result reminiscent of the multi-peaked distribution when relying on the dynamic Roy-model. Hence, search for better matches reduces the initial impact of comparative advantages by giving individuals who first sort into the bad occupation to move into high paying jobs.

For policy analysis it is crucial to investigate why the Roy-model, nested within the full model with frictions, grossly overestimates the role of type heterogeneity. Appendix Figure 11 and Figure 19 help to answer this question. The Roy-model generates type-specific life-cycle earnings distributions that hardly overlap. Their means differ substantially, while the variation around each mean is very small. Since post-labor market entry the only source of residual earnings variation, which is known to be large, are iid-shocks, the estimates sort individual into type classes which are strictly ranked by their earnings. Workers with low wealth are deemed to be of the low-type, while workers with high wealth are allocated to the high-type. Intuitively, the Roy-model interprets any unsuccessful careers as resulting from a bad set of comparative and absolute advantages, even though the true causes might be a series of bad shocks, both in terms of match quality and general skills, that accumulate over the career and are not directly related to type heterogeneity. Consequently, individuals who enter the labor market with very similar credentials but find matches of different quality and are hit by different permanent shocks are erroneously sorted into different type groups by the Roy-model.

The Roy-model also overestimates the importance of transitory skill shocks. While its exclusion from this model-specification would reduce inequality by almost 10 percent, it does not generate any inequality in the full model. Given that transitory shocks by their very nature average to zero over time, the latter result is not surprising. Its large quantitative impact in the Roy-model is due to the aforementioned dual role of targeting residual earnings variation and mobility. Since mobility decisions can have long-run career impacts, transitory shocks that influence career decisions can have permanent effects as well. The introduction of match effects and a random walk leaves little residual variation and thus a limited role for transitory skill shocks in allocating individuals across occupations.

To conclude the discussion of life-cycle inequality it is also interesting to compare the results from counterfactual exercises in the unit roots model as displayed in the third column of Table 4 to those from the full model. As discussed above, single-equation models do not control for potential selection biases. On the one hand, not controlling for the jump-process generated from search for better matches leads to an overbias of the random walk parameter because updates in match quality are mistakenly interpreted as permanent exogenous risk. On the other hand, not accounting for the mobility across employment and unemployment leads to a downward bias of the random walk parameter because particularly bad shocks pull labor market earnints below the unemployment benefits level and therefore are unobserved.

Eliminating permanent shocks reduces wealth inequality by 7 percent in the full model and by 13 percent in the single-equation unit roots model. Hence, the overbias from not modeling occupational choices outweigh the underbias from not modeling transitions between employment and unemployment. These results suggest that a considerable fraction of shocks commonly interpreted as permanent exogenous risk are in fact the outcome of a search process and reflect endogenous behavioral responses. The unit-roots model therefore over-emphasizes the importance of permanent exogenous shocks to general skills and potentially leads to a misrepresentation as to the effectiveness of social insurance to enhance economic welfare. Structural models of career dynamics are potentially very useful to explore the true impacts of social insurance systems on long-run career outcomes and wealth inequality.

6.2. The Role of Mobility

The large dispersion of match quality within occupational classes and among workers with the same skill set suggests that vertical occupational mobility and mobility between employment and unemployment are potential catalysts of inequality. To further investigate this point I simulate data for the full model but with workers not allowed to switch occupations or employment status after labor market entry. To avoid that results reflect the overprediction of unemployment rates in the first period I let individuals reconsider their choice once more at the beginning of the second year of their career. Afterward, they need to remain in that choice. Results are provided in table 5. As shown in the first row, wealth inequality would be 28 percent higher in a counterfactual world without mobility. Thus, mobility reduces rather than increases inequality. Appendix Figure 12 visualizes the effect using kernel densities of the wealth distribution. Compared to the unrestricted version if the model, there is a large mass of individuals who are permanently stuck at an extremely low earnings level, represented by workers who are still unemployed at the beginning of year 2, have initially found a very bad match or belong to types who have a comparative advantage in occupations with low returns to experience.

The rest of table 5 and Appendix Figure 13 document results from counterfactual exercises in the one-shot version of the model. As the figure clarifies, none of the exclusions completely eliminate the mass of extremely poor individuals, suggesting that they are mainly composed of permanently unemployed individuals. The exclusion of type heterogeneity, human capital or match heterogeneity have approximately the same quantitative implications on life-cycle inequality, ranging from a reduction of 26 percent for the latter to a reduction of 28 percent for the former. Ruling out human capital accumulation radically changes the shape of the wealth distribution, significantly decreasing its mean and shifting it towards the mass of extremely poor individuals. Wealth dispersion, now predominantly driven by type and match heterogeneity and permanent exogenous shocks, is still large.

I also use the one-shot version of the model to address potential concerns about the biasedness of the estimated match heterogeneity. This concern is potentially justified by the impossibility to separately identify differences in innate comparative and absolute advantages. As a consequence, match heterogeneity might reflect differences in permanent unobserved innate abilities not controlled for in the model. To provide evidence against this hypothesis I exploit the fact that models with match heterogeneity and models with uncontrolled heterogeneity in skills have very different predictions about long-run career outcomes. In particular, while the former generates a process during which workers who initially earn very little catch up to their peers, the latter implies that they remain at the bottom of the wage distribution for the rest of their career. In the one-shot model, match heterogeneity and uncontrolled permanent skill differences are indistinguishable. Thus, the difference between the wealth distribution from the one-shot model and the model allowing for mobility helps to infer the equalizing effect of mobility that would not exist when initial match heterogeneity would capture skill differences only. Given the elimination of a group of permanently low-earnings present in Appendix Figure 13 when introducing mobility speaks against this hypothesis.

6.3. Sources of Earnings Mobility

Wealth inequality is intrinsically related to earnings mobility. With perfect earnings mobility individuals would constantly change their relative position in the earnings distribution, thus eliminating any differences in life-cycle earnings. The incidence and the causes of poverty traps, a high probability that individuals who are poor today remain so in the future, are of particular interest. Given the models' multiple sources of state dependence, it naturally lends itself to the analysis of earnings mobility. In table 6 I show effects from counterfactual exercises on one year transition matrices. Both, the exclusion of type and match heterogeneity significantly alters the structure of the one-step transition matrices. Exclusion of the former reduces the diagonal elements - the probability that the individual remains within the same position of experience-specific earnings distributions - on average by seven percent, while exclusion of the latter reduces it on average by 5 percent. The effects are asymmetric. Without type heterogeneity, the likelihood of extreme transitions become more likely. For example, while the probability of being in the highest quintile in period t and being in quintile 2 in period $t+1$ and vice versa is only 0.004 in the unrestricted model, these probabilities rise to 0.13 and 0.14 respectively. The primary reason is that demotion shocks hitting workers with very high earnings periodically are not insured against by a high earnings potential anymore.

In contrast, ruling out match heterogeneity increases the mobility between Quintiles 1 and 2, and between Quintiles 4 and 5, thus making small transitions more likely. The constancy of match quality while remaining in a particular occupation locks individuals into their position within the earnings distribution. Therefore, the absence of match effects eliminates any "discreteness" of wages at the time of mobility and makes smooth transitions more likely.

These results suggest that if extreme transitions are deemed undesirable, policies reducing search frictions and helping individuals to sort into their best matches are prefered to policies that target pre-labor market skills.

7. Conclusions

In this paper I have formulated and estimated a comprehensive empirical model of life-cycle earnings dynamics and vertical occupational mobility. To obtain reliable and precise parameter estimates I have taken advantage of a unique data set from Germany that follows 56,000 employees from the time of labor market entry until twenty-three years into their careers. In a series of counterfactual experiments I have quantified the impact of pre-labor market skills, match heterogeneity, search frictions and permanent skill shocks on life-cycle earnings inequality and earnings mobility. I have found that differences in both, innate abilities and match quality, are particularly important determinants of career progression. Eliminating heterogeneity in pre-labor market skills or match quality decreases life-cycle earnings inequality by 34 and 27 percent, respectively. Furthermore, 41 percent of the variation in life-cycle earnings is observed among individuals with the same comparative advantages. Conclusions drawn from the analysis of income mobility are similar. In particular, in the absence of differences in compartive advantage or match qualities, earnings mobility increases considerably, and poverty traps become less likely.

When estimating a proto-typical Roy-model without search frictions and match heterogeneity instead - a specification dominating the literature on Discrete Choice Models of post-graduation labor market outcomes - I find the within-type variation to increase to 91 percent. This striking result is almost identical to what other work has found in US Panel Data. Thus, a model that does not control for unobserved sources of career heterogeneity that accumulate over a life-cycle erroneously interprets a large part of earnings inequality as differences in innate skills. I conclude that active labor market policies are predicted to be significantly more effective, and policies that foster per-labor market skills are likely to be significantly less effective than what is implied by earlier findings from more restrictive Roy models. Labor market policies that help individuals to find their best match and that reduce search frictions can significantly influence long-run career outcomes and reduce the incidence of poverty traps.

The model can address a broad set of research topics currently debated in areas such as

Labor Economics, Macroeconomics and Public Economics. First, since aggregate fluctuations enter the occupation-specific skill price functions, long-run career effects and welfare costs of business cycles can be quantified. Second, the framework, by explicitely solving an individual's decision problem at any point of his career, is very well suited for Policy Analysis. In contrast to work in the treatment effects literature that relies on exogenous variation at a certain point in time, one can easily simulate the long-run career effects of labor market policies. This applies even to policies that have never been established before. Third, although I have focussed on career outcomes of one education group only, the model can be extended to incorporate an education decision, therefore providing a structural framework to estimate the returns to education or to vocational training. Finally, the model, decision-theoretic in nature, can be integrated into a standard life-cycle consumption-savings model to create a unified model of career progression and wealth inequality.

References

[1] Abowd, J. M. and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*, vol. 57(2), pp. 411-445.

[2] Abowd J., F. Kramarz and D. Margolis (1999): "High-Wage Workers and High-Wage Firms." *Econometrica*, 67(2), pp. 251-333.

[3] Adda, J., C. Dustmann, C. Meghir and J.-M. Robin (2009): "Career Progression and Formal versus On-the-Job Training." Mimeo, University College London.

[4] Adda, J., M. Costa Dias, C. Meghir and B. Sianesi (2009): "Labor Market Programmes and Labor Market Outcomes." Mimeo, University College London.

[5] Aguirregabiria, V. and P. Mira (2002): "Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models." *Econometrica*, vol. 70(4), pp. 1519–1543.

[6] Aguirregabiria, V. and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games." *Econometrica*, vol. 75(1), pp. 1-53.

[7] Aguirregabiria, V. and P. Mira (2009): "Dynamic Discrete Choice Structural Models: A Survey." *Journal of Econometrics*, forthcoming.

[8] Altonji, J. G., A. A. Smith Jr. and I. Vidangos (2009): "Modeling Earnings Dynamics", NBER Working Papers w14743.

[9] Bagger, J., F. Fontaine, F. Postel-Vinay and J.-M. Robin (2006): "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation." Mimeo, University of Bristol.

[10] Baker, M. (1997): "Growth Rate Heterogeneity and the Covariance Structure of Life Cycle Earnings", *Journal of Labor Economics*, 15, pp. 338-375.

[11] Baker, M., and G. Solon (2003): "Earnings Dynamics and Inequality Among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records", *Journal of Labor Economics*, 21, pp. 289-321.

[12] Baker, M. and A. Melino (2000): "Duration Dependence and Nonparametric Heterogeneity: A Monte Carlo Study", *Journal of Econometrics*, vol. 96(2), pp. 357–393.

[13] Blundell, R., L. Pistaferri and I. Preston (2009): "Consumption Inequality and Partial Insurance", *American Economic Review*, 98(5), pp. 1887-1921.

[14] Bontemps, C., J.-M. Robin and G. J. van den Berg (1999): "An empirical equilibrium job search model with search on the job and heterogeneous workers and firms." *International Economic Review*, 40, pp. 1039-1074.

[15] Büttner, T., P. Jacobebbinghaus and J. Ludsteck (2008): "Occupational Upgrading and the Business Cycle in West Germany." Economics Discussion Papers, No 2009-34.

[16] Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica*, vol. 74(2), pp. 323-364.

[17] Cameron, A. C. and P. K. Trivedi (2005): "Microeconometrics." Cambridge University Press.

[18] Caucutt, Elizabeth and Lance Lochner (2008): "Early and Late Human Capital Investments, Borrowing Constraints, and the Family." Mimeo. University of Western Ontario.

[19] Cunha, F. (2007): "A Time to Plant and a Time to Reap." Mimeo. University of Chicago.

[20] Cunha, F., J. J. Heckman (2007): "The Technology of Skill Formation." *American Economic Review*, vol. 97(2), pp. 31-47.

[21] Cunha, F., J. J. Heckman (2007): "The Evolution of Inequality, Heterogeneity and Uncertainty in Labor Earnings in the U.S. economy." NBERworking papers, w13526.

[22] Cunha, F., J. J. Heckman (2008): “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formulation.” *Journal of Human Resources*, vol. 43(4).

[23] Devereux, P. (2002): “Occupational Upgrading and the Business Cycle.” *Labour*, vol. 16(3), pp. 423-452.

[24] Drews, N. (2007): ”Variablen der schwach anonymisierten Version der IAB-Beschaeftigten-Stichprobe 1975-2004”, FDZ Datenreport 03/2007, Nuernberg.

[25] Eckstein, Z. and K. Wolpin (1990): “Estimating a Market Equilibrium Search Model from Panel Data on Individuals.” *Econometrica*, vol. 58(4), pages 783-808.

[26] Fitzenberger, B., A. Osikumino and R. Voelter (2007): ”Imputation Rules to improve the EducationVariable in the IAB Employment Subsample”, Schmollers Jahrbuch (Journal of the Applied Social Sciences), forthcoming.

[27] Fitzenberger, B. and R. Wilke (2004): “Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System during the 1980s.” ZEW Discussion Paper No. 04-24.

[28] Floden, M. (2008): “A Note on the Accuracy of Markov-Chain Approximations to Highly Persistent AR(1) Processes.” *Economics Letters*, 99, pp. 516 - 520.

[29] Geweke, J. (1989): ”Bayesian Inference in Econometric Models Using Monte-Carlo Integration.” *Econometrica*, 57, pp. 1317-1339.

[30] Gibbons, R. and M. Waldman (1999): “A Theory of Wage and Promotion Dynamics inside Firms.” *Quarterly Journal of Economics*, vol. 114(4), pp. 1321-1358.

[31] Gibbons, R. and M. Waldman (2006): “Enriching a Theory of Wage and Promotion Dynamics inside Firms.” *Journal of Labor Economics*, vol. 24(1), pp. 59-108.

[32] Gottschalk, P., and R. A. Moffitt (1994): ”The Growth of Earnings Instability in the U.S. Labor Market”, *Brookings Paper on Economic Activity*, 25(2), pp. 217-272.

[33] Gourieroux, C. and A. Monfort (1996): "Simulation-Based Econometric Methods." CORE Lectures, Oxford University Press.

[34] Gourinchas, P.-O., and J. A. Parker (2001): "The Empirical Importance of Precautionary Saving", *American Economic Review*, 91(2), pp. 406-412.

[35] Gourinchas, P.-O., and J. A. Parker (2002): "Consumption over the Life-Cycle", *Econometrica*, 70(1), pp. 47-89.

[36] Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks really very Persistent?", *American Economic Review*, 97(3), pp. 687-712.

[37] Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes", *Review of Economic Dynamics*, 12(1), pp. 58-79.

[38] Guvenen, F. and A. A. Smith Jr. (2009): "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach." mimeo, University of Minnesota.

[39] Haider, S. (2001): "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991", *Journal of Labor Economics*, 19(4), pp. 799-836.

[40] Hajivassiliou, V. and D. McFadden (1998): "The Method of Simulated Scores for the Estimation of LDV models." *Econometrica*, 66, pp. 863-896.

[41] Hall, R. E., and F. S. Mishkin (1982): "The Sensitivity of Consumption to Transitory Income: Estimation from Panel Data on Households", *Econometrica*, 50(2), pp. 461-481.

[42] Hause, J. (1980): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis." *Econometrica*, vol. 48(4), pp. 1013 – 1029.

[43] Heathcote, J., K. Storesletten, and G. L. Violante (2005): "Two Views of Inequality over the Life Cycle", *Journal of the European Economic Association*, 3, pp. 765-775.

[44] Heathcote, J., K. Storesletten, and G. L. Violante (2008): "Insurance and Opportunities: A Welfare Analysis of Labor Market Risk", *Journal of Monetary Economics*, 55(3), pp. 501-525.

[45] Heckman, J. J. and B. Singer (1984): "A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models of Durations Data." *Econometrica*, vol. 52(2), pp. 271-320.

[46] Heckman, J. J., J. Stixrud and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, vol. 24(3), pp. 411-482.

[47] Hoffmann, F. (2009): "The Different Life-Cycle Income Trajectories of Less and More Educated Workers." mimeo, University of Toronto.

[48] Hornstein, A., P. Krusell and G. Violante (2006): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." CEPR Discussion Paper 5937.

[49] Hotz, V. J. and R. A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models." *The Review of Economic Studies*, 60(3), pp. 497–529.

[50] Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994): "A simulation estimator for dynamic models of discrete choice." *The Review of Economic Studies*, 61(2), pp. 265–289.

[51] Huggett, M. and G. Ventura (1999): "On the Distributional Effects of Social Security Reform." *Review of Economic Dynamics*, vol. 2(3), pages 498-531.

[52] Huggett, M., G. Ventura, and A. Yaron (2006): "Human Capital and Earnings Distribution Dynamics", *Journal of Monetary Economics*, 53(2), pp. 265-290.

[53] Huggett, M., G. Ventura, and A. Yaron (2007): "Sources of Lifetime Inequality", NBER Working Papers 13224.

[54] Hunt, J. (1995): “The Effect of Unemployment Compensation on Unemployment Duration in Germany.” *Journal of Labor Economics*, vol. 13(1), pp. 88-120.

[55] Imai, S., N. Jain, and A. Ching (2006): “Bayesian estimation of dynamic discrete choice models.” mimeo, Rotman School of Business at the University of Toronto.

[56] Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): “The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the U.S.” *European Economic Review*, vol. 50(4), pp. 877-907.

[57] Jovanovic, B. (1979): ”Job Matching and the Theory of Turnover.” *Journal of Political Economy*, 87(5), pp. 972-90.

[58] Judd, K. L. (1998): “Numerical Methods in Economics.” The MIT Press, Cambridge, Massachusetts.

[59] Kambourov, G. and I. Manovskii (2009): “Occupational Mobility and Wage Inequality.” *Review of Economic Studies*, vol. 76(2), pp. 731-759.

[60] Kaplan, G. (2007): ”Inequality and the Life-Cycle.” Mimeo, University of Pennsylvania.

[61] Keane, M. (1994): ”A Computationally Practical Simulation Estimator for Panel Data.” *Econometrica*, 62, pp. 95-116.

[62] Keane, M. P., and K. I. Wolpin (1997): ”The Career Decisions of Young Men”, *Journal of Political Economy*, 105(3), pp. 473-522.

[63] Krueger, D., and F. Perri (2006): ”Does Income Inequality Lead to Consumption Inequality? Evidence and Theory”, *Review of Economic Studies*, 73(1), pp. 163-193.

[64] Lentz, Rasmus (2009): “Sorting by Search Intensity.” *Journal of Economic Theory*, forthcoming.

[65] Lise, J. (2007): “On-the-Job Search and Precautionary Savings: Theory and Empirics of Earnings and Wealth Inequality.” Mimeo, University College London.

[66] Lopes de Melo, Rafael (2008): "Sorting in the Labor Market: Theory and Measurement." Mimeo, Yale University.

[67] Low, H., C. Meghir and L. Pistaferri (2008): "Wage Risk and Employment Risk over the Life Cycle", IFS Working Papers W08/06.

[68] MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." *Journal of Econometrics*, vol. 18, pp. 83-114.

[69] Mariano, R, T. Schuermann and M. J. Weeks (2000): "Simulation-Based Inference in Econometrics." Cambridge University Press.

[70] Menzio, G and S. Shi (2008): "Efficient Search On-the-Job and the Business Cycle." Mimeo, University of Toronto.

[71] Menzio, G and S. Shi (2009): "Efficient Search On-the-Job and the Business Cycle." *Journal of Economic Theory*, forthcoming.

[72] Miller, R. (1984): "Job matching and Occupational Choice." *Journal of Political Economy*, 92(6), pp. 1086-1120.

[73] Mincer, J. (1974): Schooling, Experience and Earnings, Columbia University Press, New York.

[74] Moffitt, R. A., and P. Gottschalk (2002): "Trends in the Transitory Variance of Earnings in the United States", *Economic Journal*, 112, pp. C68-C73.

[75] Moscarini, G and F. Postel-Vinay (2009): "Non-stationary Search Equilibrium." Mimeo, Yale University.

[76] Neal, D. (1999): "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics*, 17(2), pp. 237-61.

[77] Norets, Andriy (2009): "Inference in Dynamic Discrete Choice Models with Serially Correlated Unobserved State Variables." *Econometrica*, vol. 77(5), pp. 1665-1682.

[78] Pavan, R. (2009): "Career Choice and Wage Growth." Mimeo, University of Rochester.

[79] Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, vol. 70(6), pp. 2295-2350.

[80] Postel-Vinay, F. and H. Turon (2009): "On-the-job Search, Productivity Shocks, and the Individual Earnings Process." *International Economic Review*, forthcoming.

[81] Quah, D. (1990): "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption", *Journal of Political Economy*, 98(3), pp. 449-475.

[82] Restuccia, D., and C. Urrutia (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education." *American Economic Review*, 95(5), pp. 1354–78.

[83] Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, Vol. 55(5), pp. 999-1033.

[84] Shi, S. (2009): "Directed Search for Equilibrium Wage-Tenure Contracts." *Econometrica*, vol. 77, pp. 561-584.

[85] Storesletten, K, C. Telmer, and A. Yaron (2004a): "Consumption and Risk-Sharing over the Life-Cycle", *Journal of Monetary Economics*, 51, pp. 609-633.

[86] Storesletten, K, C. Telmer, and A. Yaron (2004b): "Cyclical Dynamics in Ideosyncratic Labor Market Risk", *Journal of Political Economy*, 112, pp. 695-717.

[87] Sullivan, P. (2009): "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search." *International Economic Review*, forthcoming.

[88] Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters*, 20, pp. 177 - 181.

[89] Tauchen, G. and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models." *Econometrica*, 59(2), pp. 371 - 396.

[90] Topel, R. H., and M. P. Ward (1992): "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 107(2), pp. 439 - 479.

[91] Train, K. (2003): "Discrete Choice Methods with Simulation." Cambridge University Press.

Appendix

A. The Likelihoods in the Full Model

An individual's likelihood contribution does not have a closed form. However, the procedure of computing the likelihood can be facilitated considerably by taking advantage of the economically plausible iid-assumption of transitory occupation-specific shocks. In a first step I condition on all variables but the transitory shocks. The conditioning variables are summarized in the state vector $\widetilde{S}_{it} = \left(X_{it}^j, k, \mu_{it}^j, u_{it}\right)_{j\in\{u,b,m,g\}}$ which varies across individuals and over time. The only remaining random element are therefore the transitory alternative-specific shocks ε_{it}^j. The joint probability of the observed wage and choice in period t for individual i is given by

$$L_{it} \equiv \Pr(w_{it}^{j^*}, j_i^*(t) \mid \widetilde{S}_{it}) = \Pr(w_{it}^{j^*(t)} \mid \widetilde{S}_{it}) * \Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}), \tag{A.1}$$

the product of the conditional wage density and the conditional choice probability. Let $\widehat{w}_{it}^j = w_{it}^j\left(\widetilde{S}_{it}\right)$ denote the wage predicted from the state variables, and let ϕ be the pdf of $N(0,1)$. Then the conditional wage density is given by a standard OLS-expression

$$\Pr(w_{it}^{j^*(t)} \mid \widetilde{S}_{it}) = \phi\left(\frac{w_{it}^{j^*(t)} - \widehat{w}_{it}^{j^*(t)}}{\sqrt{var\left(w_{it}^{j^*(t)} - \widehat{w}_{it}^{j^*(t)}\right)}}\right). \tag{A.2}$$

This term is unaffected by introducing search frictions or Dynamic Programming.

Computing the conditional choice probabilities is more involved. Due to market frictions, they depend on the previous' period choice. Let $V_{it}^{j}\left(w_{it}^{j}, \widetilde{S}_{it}\right)$ be the alternative-specific value functions, as described by equations (3.11) and (**??**). To compute choice probabilities we need to consider the following cases:

A.0.1. The worker is observed to be employed ($j^*(t) \neq u$), and it is not his year of labor market entry:

- The individual is observed in the same occupation as in the previous period:

$$\Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}) = (1-\delta-\kappa)*\left[\begin{array}{c} (1-\lambda_1)*\Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{u} \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}\right) \\ +\lambda_1 * \Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{k}; \forall k \neq j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}\right) \end{array}\right]. \quad \text{(A.3)}$$

- The individual is observed in a different occupation as in the previous period:

$$\Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}) = (1-\delta)*\left[\begin{array}{c} \left(\frac{\lambda_1}{\lambda_1+\kappa}\right)*\Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{k}; \forall k \neq j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}\right) \\ +\left(\frac{\kappa}{\lambda_1+\kappa}\right)*\Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{k}; \forall k \neq j_i^*(t-1)\right) \end{array}\right]. \quad \text{(A.4)}$$

- The individual was unemployed in the previous period:

$$\Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}) = \lambda_0 * \Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{k}; \forall k \neq j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}\right). \quad \text{(A.5)}$$

A.0.2. The worker is observed to be employed ($j^*(t) \neq u$), and it is his year of labor market entry:

$$\Pr(j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}) = \lambda_0 * \Pr\left(V_{it}^{j_i^*(t)} > V_{it}^{k}; \forall k \neq j_i^*(t) \mid w_{it}^{j^*(t)}, \widetilde{S}_{it}\right). \quad \text{(A.6)}$$

A.0.3. The worker is observed to be unemployed ($j^*(t) = u$):

- The individual was employed in the previous period:

$$\Pr(u \mid b^u_{it}, \widetilde{S}_{it}) = \delta + (1-\delta) * \left[\begin{array}{c} (1-\lambda_1-\kappa) * \Pr\left(V^u_{it} > V^{j^*_i(t-1)}_{it} \mid b^u_{it}, \widetilde{S}_{it}\right) \\ +\lambda_1 * \Pr\left(V^u_{it} > V^k_{it}; \forall k \neq u \mid b^u_{it}, \widetilde{S}_{it}\right) \\ +\kappa * \Pr\left(V^u_{it} > V^k_{it}; \forall k \neq u, j^*_i(t-1) \mid b^u_{it}, \widetilde{S}_{it}\right) \end{array} \right]. \quad \text{(A.7)}$$

- The individual was unemployed last period:

$$\Pr(u \mid b^u_{it}, \widetilde{S}_{it}) = (1-\lambda_0) + \lambda_0 * \Pr\left(V^u_{it} > V^k_{it}; \forall k \neq u \mid b^u_{it}, \widetilde{S}_{it}\right). \quad \text{(A.8)}$$

B. Simulation of the Bellman-Equations and Construction of the Likelihood

To compute the likelihoods I use a nested algorithm that simulates the Bellman-equation in the inner step and that simulates the likelihoods in the outer step. The resulting simulated likelihood is maximized using the algorithm proposed by Broyden, Fletcher, Goldfarb and Shanno, a gradient-based method that has been found to have desirable numerical properties in the estimation of Discrete Choice Models.

Let T_i denote the length of the career trajectory observed for individual i, and let R denote the number of histories of length T_i drawn for each individuals. The algorithm is described as follows:

1. For each individual in the data, draw R histories with length T_i of $N(0,1)$ variables for the random-walk and the three processes describing the evolution of the match effects. These draws need to be used in every iteration of the algorithm. I use a different set of draws for each type. With four types and four dynamic unobserved processes, there are in total $4 * 4 * R * T_i$ draws per individual.

2. Choose a starting value for the parameters. In the estimation of the Dynamic Programming framework I use the estimates obtained from the Dynamic Probit models.

3. Construct the random walk and the three processes of match effects using the draws in Step 1 and the parameters in step 2.

4. Given the current value of the parameter vector, simulate the Dynamic Programming Problem by backward recursion. The state-space includes observables, the state of the random walk and the state of the match effects. Conditional on these state variables, the transitory shocks can be integrated numerically.

5. Match the Bellman-equations to the data observed for individual i. Random draws are treated as if observable. Since the state-space of the Bellman-equations is discretized, while the random draws are continous, I use a nearest neighborhood matching procedure. To account for the continuity of the draws, Bellman-equations are then interpolated between the discretized state-points.

6. Construct the experience$-t$, individual$-i$ likelihoods conditional on the observables, the random walk and the process of match effects. From this, construct the individual$-i$ likelihood. There are R such expressions. To compute the empirical likelihood (4.3), take an average over these R expressions.

7. If the change of the likelihood across two sub-sequent iterations is larger than a certain threshold criterion, update the parameter vector using the BFGS-expression and go back to Step 2. Otherwise use the current parameter vector as estimates.

As discussed in section 4.2, the simulated likelihood is continuous. Consequently, fast gradient-based numerical maximization algorithms such as BFGS work well.

In regards to the simulation of the Dynamic Programming problem, two comments are in order. First, to discretize the states of the random walk and the match effects, I use Tauchen's (1986) method.[38] For the match effects, whose distribution is stationary, this procedure is straightforward. However, the random walk is non-stationary. I apply Tauchen's method for each pair of sub-sequent periods separately and allow the state-space of the random-walk to grow at a linear rate.

Second, to interpolate the value function between discretized state-points, I use a multi-dimensional simplex-method. This algorithm, as described in Judd (1998), p.243, is faster and more reliable than linear multidimensional interpolation.

[38] Tauchen and Hussey's (1991) method is an alternative. I have chosen Tauchen's (1986) method instead because this algorithm has been found to be more robust than Tauchen and Hussey's procedure. See for example Floden (2008).

TABLE 1: SAMPLE SIZE BY EXPERIENCE

Experience (Potential)	Number of Observations	
	Full Sample	10% Sub-Sample
0	55,677	5,592
1	55,677	5,592
2	55,677	5,592
3	55,677	5,592
4	55,677	5,592
5	55,003	5,520
6	52,561	5,272
7	50,325	5,054
8	48,254	4,854
9	46,228	4,662
10	42,872	4,300
11	39,543	3,969
12	36,313	3,610
13	33,064	3,274
14	29,903	2,968
15	26,859	2,670
16	24,030	2,390
17	21,146	2,109
18	18,419	1,839
19	15,748	1,567
20	13,248	1,324
21	10,842	1,104
22	8,632	873
TOTAL	851,375	85,319

NOTES: The 10 percent sub-sample is constructed from a random draw of employees from the full sample. Once an individual is chosen, his whole labor market career is kept in the sample. For computational reasons, the 10 percent sub-sample is used in the estimation.

TABLE 2: FRACTION OF OCCUPATIONAL MOBILITY ASSOCIATED WITH WAGE RISES/FALLS

		Wage Increase		Wage Decrease	
		Full Sample	10% Sample	Full Sample	10% Sample
Upward Mobility	Bad to Mediocre	0.46	0.47	0.37	0.35
	Bad to Good	0.59	0.63	0.26	0.22
	Mediocre to Good	0.60	0.61	0.25	0.24
Downward Mobility	Good to Mediocre	0.50	0.49	0.33	0.32
	Good to Bad	0.51	0.58	0.30	0.26
	Mediocre to Bad	0.48	0.49	0.35	0.35

NOTES: This table shows the fraction of occupational changes that are associated with discrete wage increases/decreases. "Upward Mobility" is defined as occupational mobility into a better occupation, and "Downward Mobility" is defined as occupational mobility into a worse occupation. The algorithm allocating 3-digit occupations into the three groups refered to as "bad", "mediocre" and "good" is described in the text. A discrete wage change is defined as a wage change that is larger than 10 percent of the standard deviation of wages in the sample.

TABLE 3: INCOME MOBILITY

Panel 1: 1-Year Transition Matrices

		Period t+1				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.798	0.140	0.033	0.018	0.011
		0.798	*0.144*	*0.030*	*0.017*	*0.011*
	Quintile 2	0.153	0.656	0.158	0.026	0.007
		0.151	*0.653*	*0.161*	*0.027*	*0.007*
Period t	Quintile 3	0.034	0.174	0.611	0.161	0.019
		0.034	*0.174*	*0.608*	*0.163*	*0.021*
	Quintile 4	0.020	0.022	0.177	0.643	0.138
		0.020	*0.021*	*0.179*	*0.642*	*0.138*
	Quintile 5	0.013	0.005	0.015	0.146	0.820
		0.013	*0.006*	*0.017*	*0.145*	*0.820*

Panel 2: 5-Year Transition Matrices

		Period t+5				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.617	0.215	0.087	0.050	0.031
		0.617	*0.222*	*0.086*	*0.043*	*0.032*
	Quintile 2	0.248	0.415	0.215	0.089	0.034
		0.247	*0.405*	*0.217*	*0.096*	*0.035*
Period t	Quintile 3	0.092	0.250	0.371	0.213	0.074
		0.089	*0.251*	*0.366*	*0.218*	*0.076*
	Quintile 4	0.048	0.090	0.248	0.404	0.210
		0.045	*0.089*	*0.252*	*0.405*	*0.207*
	Quintile 5	0.026	0.025	0.067	0.234	0.649
		0.025	*0.028*	*0.069*	*0.230*	*0.648*

Panel 3: 10-Year Transition Matrices

		Period t+10				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.504	0.238	0.127	0.080	0.050
		0.494	*0.249*	*0.130*	*0.074*	*0.053*
	Quintile 2	0.255	0.319	0.226	0.135	0.065
		0.255	*0.312*	*0.227*	*0.141*	*0.064*
Period t	Quintile 3	0.131	0.237	0.286	0.224	0.122
		0.131	*0.239*	*0.282*	*0.228*	*0.120*
	Quintile 4	0.073	0.133	0.235	0.314	0.245
		0.073	*0.126*	*0.237*	*0.314*	*0.249*
	Quintile 5	0.038	0.051	0.109	0.249	0.552
		0.035	*0.051*	*0.111*	*0.253*	*0.551*

NOTES: This table shows the probability that a worker whose income falls into the q-th quintile of experience-specific income distributions receives income in the p-th quintile 1 year, 5 years and 10 years later. Wages are corrected from cohort effects. Values from the 10 percent sub-sample are in *Italic* .

TABLE 4: THE ROLE OF DIFFERENT VARIANCE COMPONENTS ON LIFE-CYCLE INEQUALITY

PANEL A: PERCENTAGE CHANGE OF THE STANDARD DEVIATION OF WEALTH IN COUNTERFACTUAL EXPERIMENTS:

	Model Specification		**Comparison**
	Multi-Period Roy Model (Keane&Wolpin)	Full Model With Frictions	Simple Random Walk model
Types	-73.8	-34.2	-61.0
Transitory Shocks	-14.0	-0.3	-0.8
Permanent Shocks	-	-7.1	-12.8
Match Effects	-	-27.3	-
Frictions	-	5.2	-
PANEL B: ACROSS TYPE VARIATION IN WEALTH:			
	91.2	40.9	72.4

NOTES: This table displays results from counterfactual experiments. Each column refers to a different specification of the Dynamic Discrete Choice Model. For comparison, results from a regression model with random walk shocks are shown as well. Each of the cells in Panel A show the percentage changes of wealth inequality, measured by its standard deviation of total life-cycle earnings, from a different counterfactual experiment. Counterfactuals are constructed as follows: The model is simulated for 10,000 individuals and 22 years using the original parameter estimates, with one set of parameter estimates per row adjusted. For example, the row "Transitory Shocks" lists the effect on wealth inequality when simulating a particular model specification using the original parameter estimates, but with the variance of transitory shocks set to zero. Panel B displays results from a simple across-type variance decomposition.

TABLE 5: MOBILITY AND INEQUALITY

PANEL A: PERCENTAGE CHANGE OF THE STANDARD DEVIATION OF WEALTH WHEN MOBILITY IS RULED OUT:

Relative to Unrestricted Full Model	28.4
Counterfactuals in Model without Mobility	
No Types	-28.2
No Initial Match Heterogeneity	-26.1
No Human Capital Accumulation	-27.4

NOTES: This table shows percentage changes of wealth inequality, measured by its standard deviation of total life-cycle earnings, when ruling out mobility. Results from counterfactual exercises in this restricted version of the model are provided as well. The simulated data are constructed as follows: The model is simulated for 10,000 individuals and 22 years using the original parameter estimates, but with mobility after labor market entry ruled out. Counterfactuals are simulated by adjusting the parameter values to a value reflecting the experiment. For example, in the counterfactual exercise "No Initial Match Heterogeneity", the variance of match effects are equal to zero.

TABLE 6: THE ROLE OF VARIANCE COMPONENTS ON INCOME MOBILITY IN THE FULL MODEL

		FULL MODEL WITH FRICTIONS					NO TYPES				
		Period t+1					Period t+1				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.647	0.218	0.063	0.039	0.033	*0.599*	*0.235*	*0.086*	*0.046*	*0.034*
	Quintile 2	0.238	0.475	0.238	0.045	0.004	*0.251*	*0.406*	*0.247*	*0.081*	*0.014*
Period t	Quintile 3	0.057	0.256	0.427	0.232	0.029	*0.081*	*0.257*	*0.355*	*0.246*	*0.061*
	Quintile 4	0.029	0.045	0.243	0.474	0.208	*0.036*	*0.086*	*0.251*	*0.401*	*0.226*
	Quintile 5	0.026	0.004	0.029	0.212	0.730	*0.028*	*0.013*	*0.061*	*0.229*	*0.670*

		NO MATCH HETEROGENEITY					NO FRICTIONS				
	Quintile 1	*0.577*	*0.256*	*0.089*	*0.042*	*0.037*	*0.704*	*0.233*	*0.052*	*0.009*	*0.002*
	Quintile 2	*0.271*	*0.425*	*0.248*	*0.051*	*0.005*	*0.244*	*0.442*	*0.242*	*0.062*	*0.010*
Period t	Quintile 3	*0.092*	*0.259*	*0.383*	*0.224*	*0.041*	*0.043*	*0.259*	*0.400*	*0.243*	*0.054*
	Quintile 4	*0.031*	*0.053*	*0.231*	*0.433*	*0.251*	*0.003*	*0.059*	*0.257*	*0.446*	*0.236*
	Quintile 5	*0.025*	*0.004*	*0.048*	*0.252*	*0.671*	*0.000*	*0.003*	*0.049*	*0.243*	*0.705*

NOTES: This table displays results from counterfactual experiments in the full model with frictions. It shows one-step transition matrices for income - the probabilities that a worker whose income falls into the q-th quintile of experience-specific income distributions receives income in the p-th quintile 1 year later - from different counterfactuals. Wages are corrected from cohort effects. Counterfactuals are constructed as follows: The model is simulated for 10,000 individuals and 22 years using the original parameter estimates, with one set of parameter estimates per matrix adjusted. For example, the matrix "No Permanent Shocks" lists the effect on income mobility when simulating a particular model specification using the original parameter estimates, but with the variance of permanent shocks set to zero.

FIGURE 1

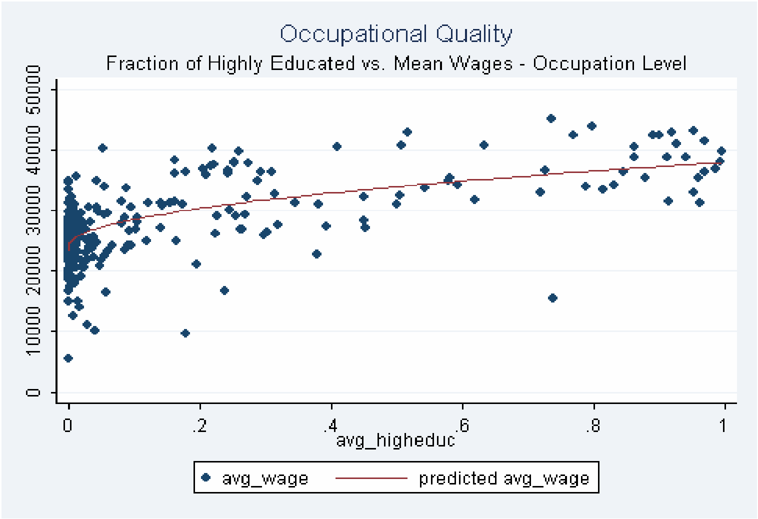


FIGURE 2

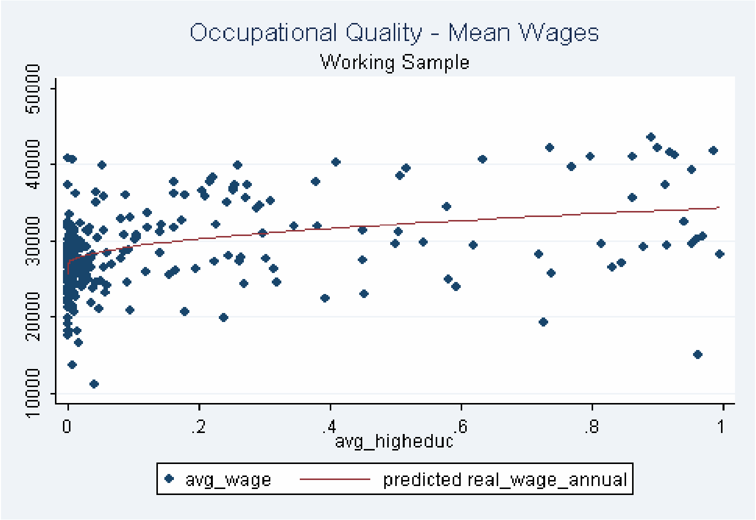


FIGURE 3

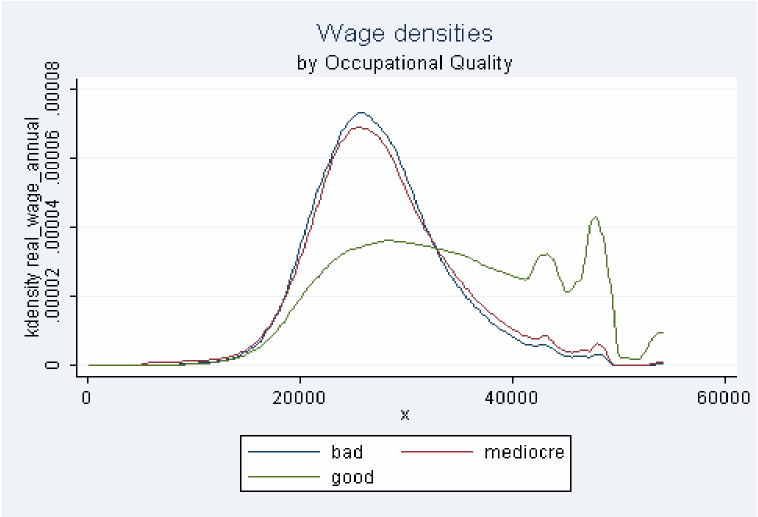


FIGURE 4

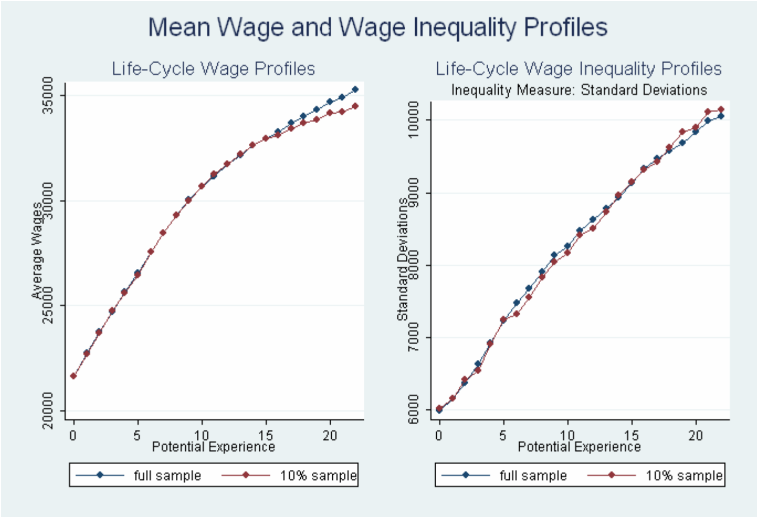


FIGURE 5

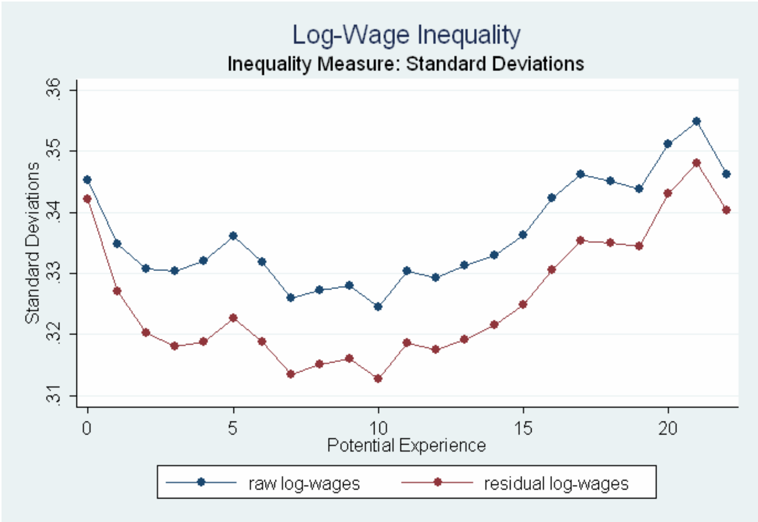


FIGURE 6

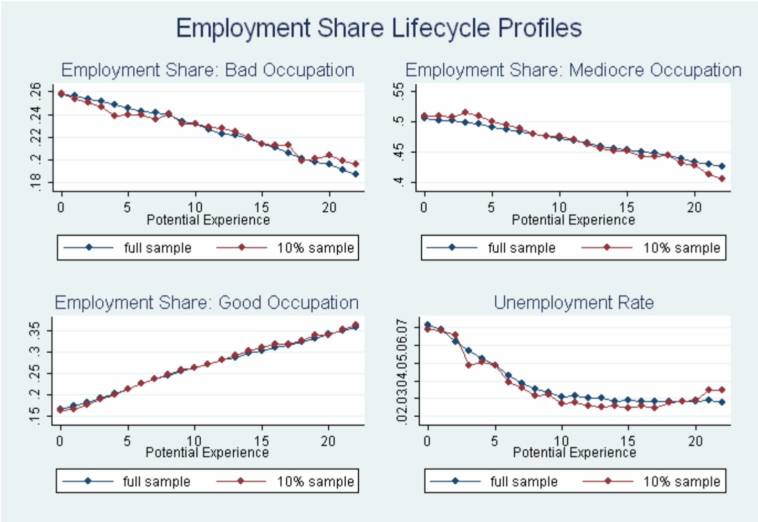


FIGURE 7

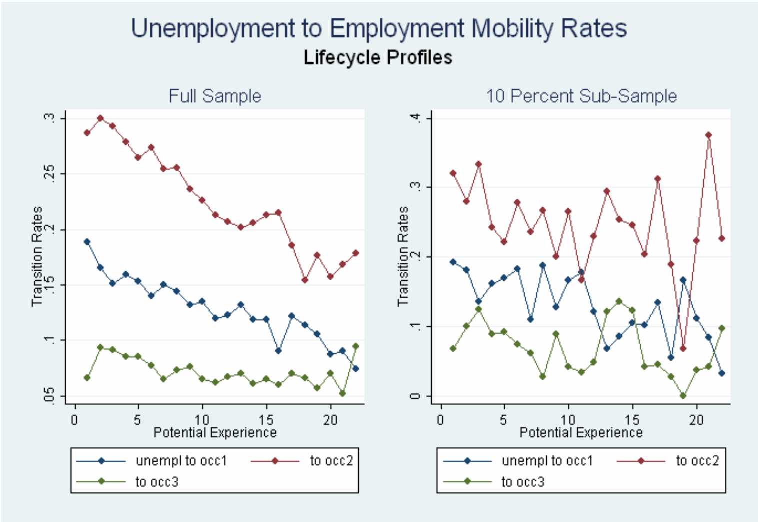


FIGURE 8

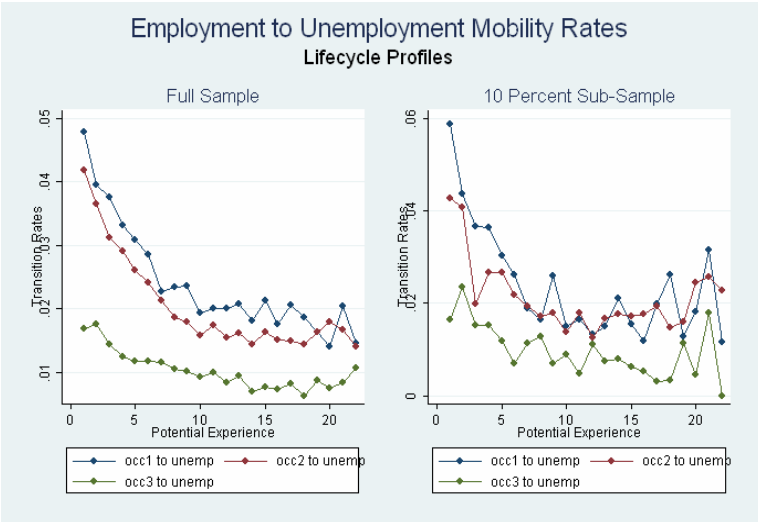


FIGURE 9

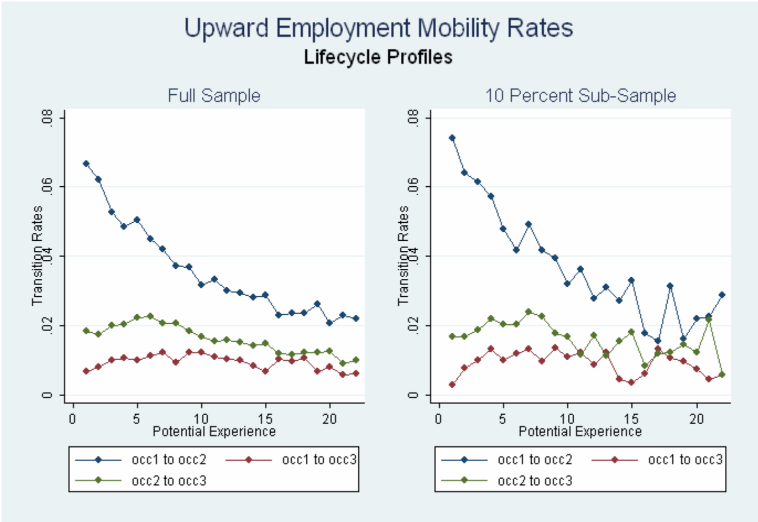


FIGURE 10

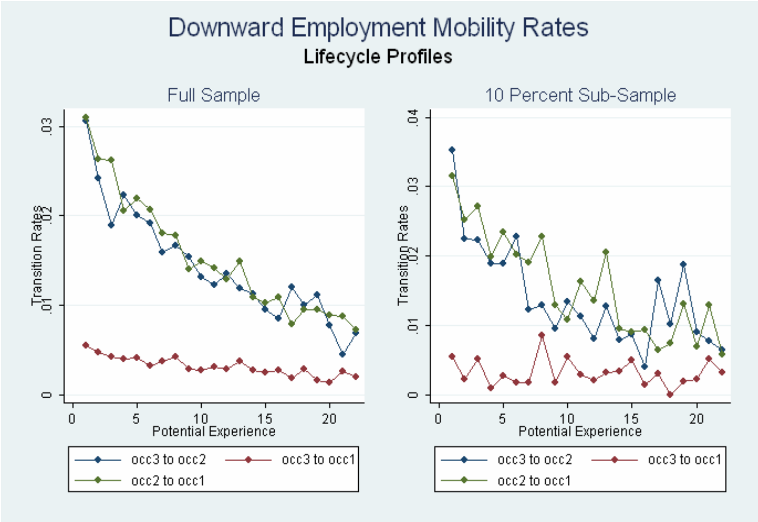


FIGURE 11

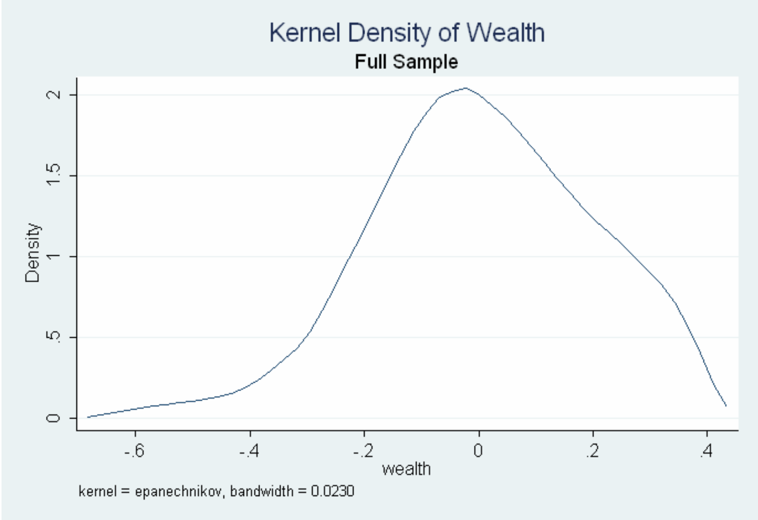


FIGURE 12

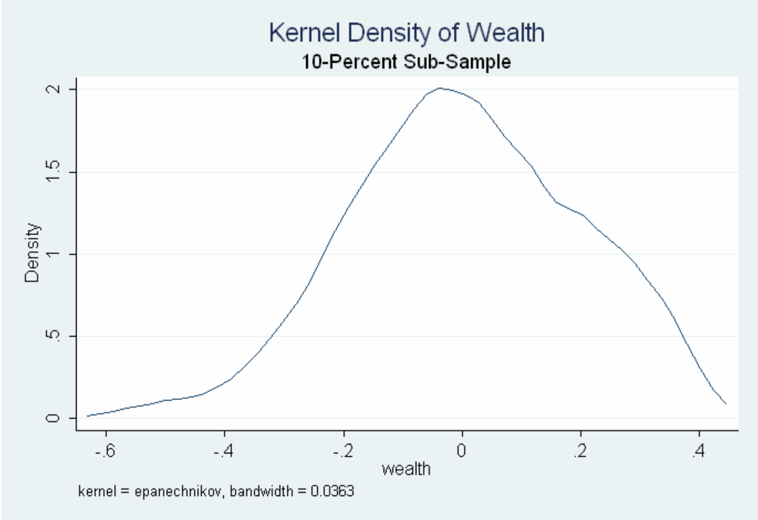


FIGURE 13 - LABOR MARKET TRANSITIONS IN THE MODEL

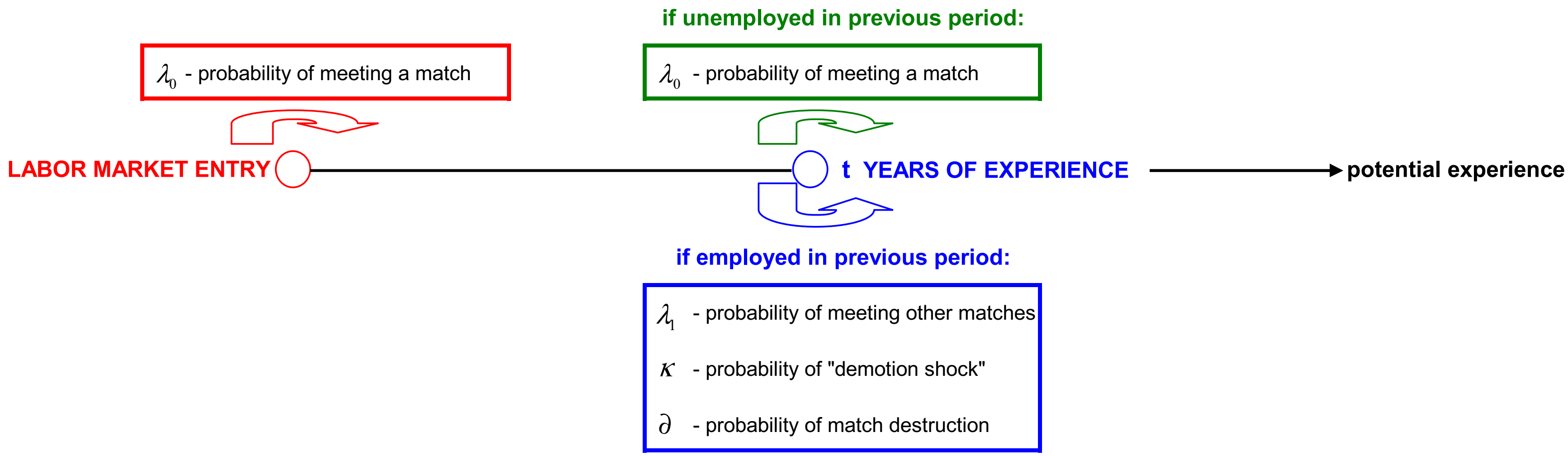


FIGURE 14

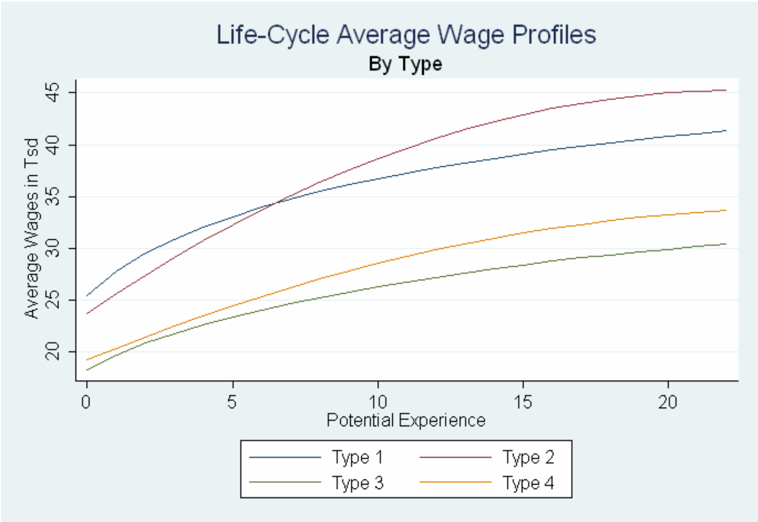


FIGURE 15

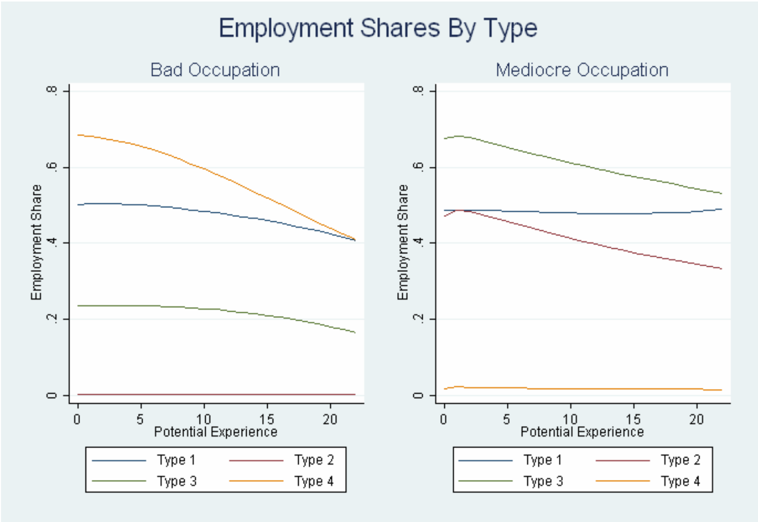


FIGURE 16

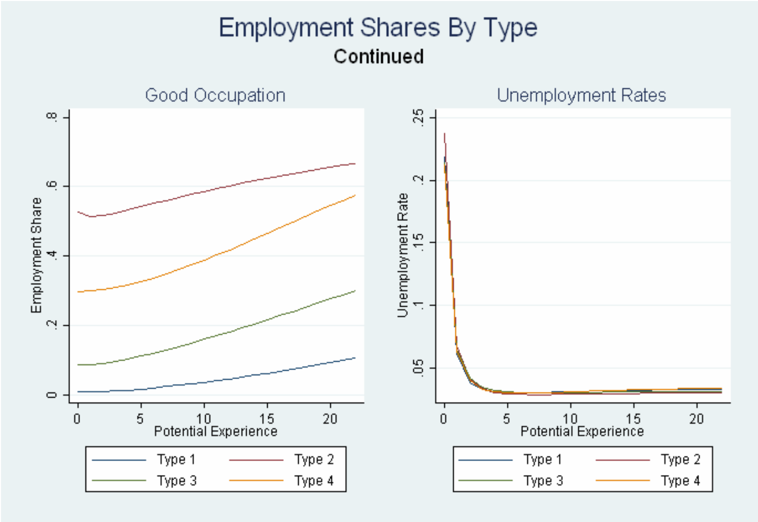


FIGURE 17

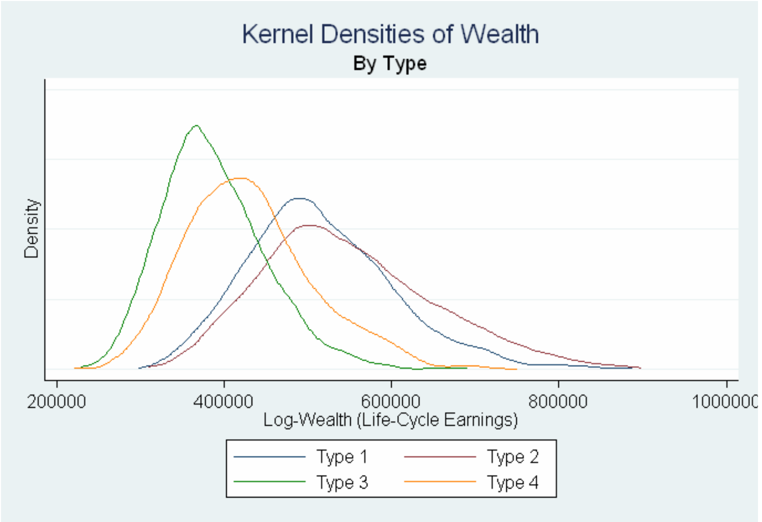


FIGURE 18

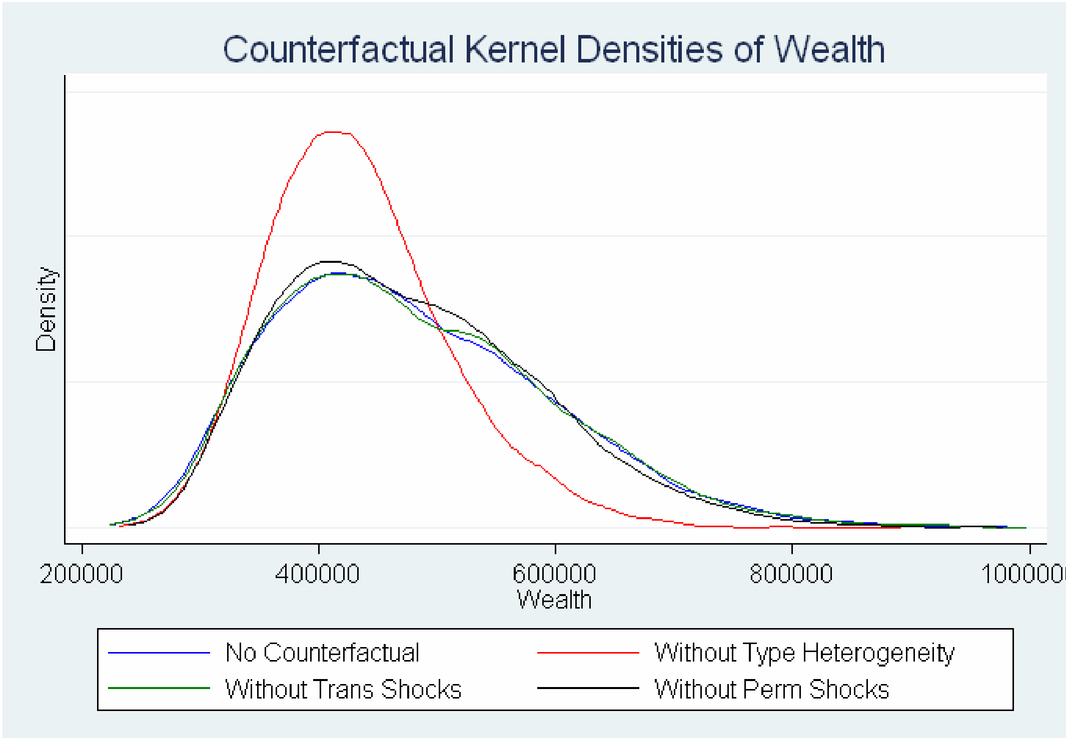
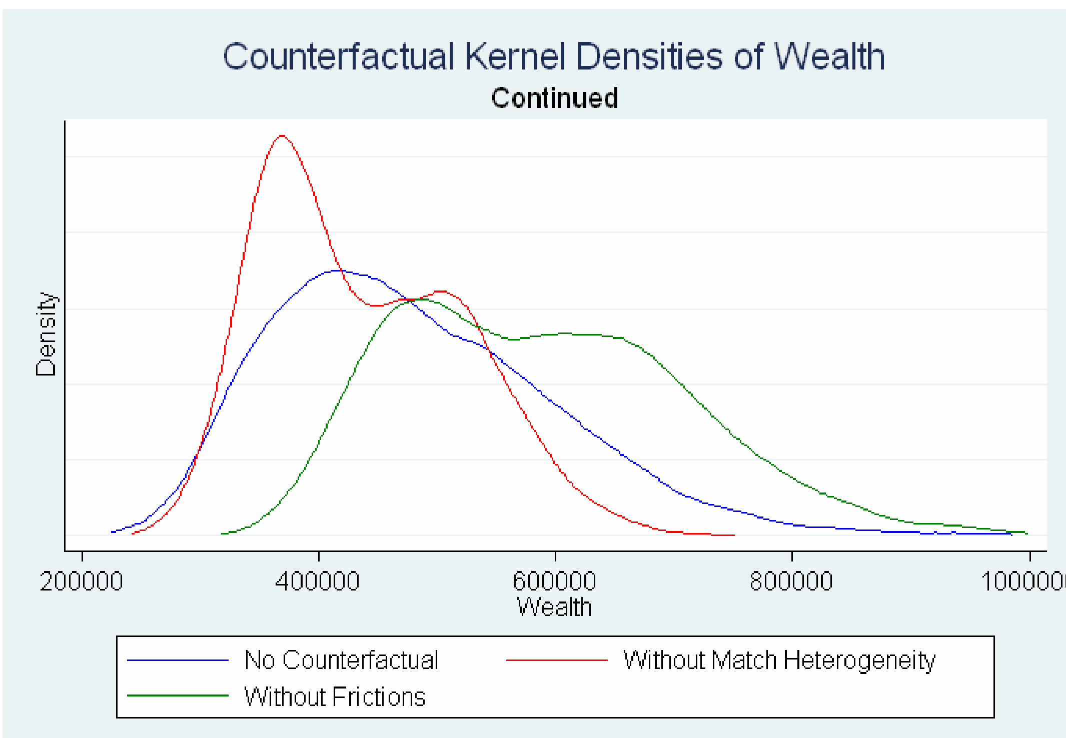


FIGURE 19



APPENDIX TABLE 1: PARAMETER ESTIMATES FOR FULL MODEL WITH FRICTIONS

PANEL A: OBSERVED HETEROGENEITY

	Intercept, Type 1	9.525	
		0.000	***
	Intercept, Type 2	8.820	
		0.000	***
	Intercept, Type 3	9.039	
		0.000	***
	Intercept, Type 4	9.228	
		0.000	***
	Time Trend	0.211	
		0.000	***
EQUATION 1 (BAD OCCUPATION)	Time Trend^2	0.077	
		0.008	***
	Unemployment Rate	-0.961	
		0.052	***
	Experience	0.028	
		0.001	***
	Experience^2	-0.001	
		0.000	***
	Tenure	0.004	
		0.001	***
	Tenure^2	0.000	
		0.000	***
	Intercept, Type 1	9.395	
		0.000	***
	Intercept, Type 2	9.281	
		0.000	***
	Intercept, Type 3	9.091	
		0.000	***
	Intercept, Type 4	8.640	
		0.000	***
	Time Trend	0.277	
		0.000	***
EQUATION 2 (MEDIOCRE OCCUPATION)	Time Trend^2	0.074	
		0.000	***
	Unemployment Rate	-0.922	
		0.037	***
	Experience	0.027	
		0.001	***
	Experience^2	0.000	
		0.000	***
	Tenure	0.008	
		0.001	***
	Tenure^2	-0.001	
		0.000	***

APPENDIX TABLE 1: CONT'D

EQUATION 3 (GOOD OCCUPATION)	Intercept, Type 1	8.712	
		0.020	***
	Intercept, Type 2	9.131	
		0.016	***
	Intercept, Type 3	8.581	
		0.018	***
	Intercept, Type 4	8.786	
		0.017	***
	Time Trend	0.376	
		0.000	***
	Time Trend^2	0.065	
		0.000	
	Unemployment Rate	-0.678	
		0.047	***
	Experience	0.068	
		0.001	***
	Experience^2	-0.002	
		0.000	***
	Tenure	0.001	
		0.000	***
	Tenure^2	0.000	
		0.000	***
EQUATION 4 (UNEMPLOYMENT BENEFITS)	Intercept	8.803	
		0.000	***
	Time Trend	0.065	
		0.058	
	Time Trend^2	0.009	
		0.000	***
	Unemployment Duration	-0.018	
		0.012	
	Unemployment Duration^2	-0.009	
		0.001	**
	max{Last_wage,0}	0.017	
		0.001	***
	max{Last_wage,0} *Unemployment Duration	0.004	
		0.001	***
	Dummy(duration>=2years)	-0.024	
		0.022	

APPENDIX TABLE 1: CONT'D

PANEL B: UNOBSERVED HETEROGENEITY

VARIANCE COMPONENTS (IN STANDARD DEVIATIONS)	Trans.Shock, Bad Occ	0.081	
		0.001	***
	Trans. Shock, Mediocre Occ	0.085	
		0.000	***
	Trans Shock, Good Occ	0.082	
		0.001	***
	Trans Shock, Unempl	0.260	
		0.003	***
	Perm Shock, General Skills	0.035	
		0.000	***
	Match Effects	0.183	
		0.001	***
TYPES AND TYPE PROPORTIONS	Fraction of Type 1	0.249	***
	Fraction of Type 2	0.247	***
	Fraction of Type 3	0.253	***
	Fraction of Type 4	0.250	***
FRICTIONS	Lambda 0	0.782	
		0.004	***
	Lambda 1	0.020	
		0.001	***
	Delta	0.025	
		0.001	***
	Kappa	0.013	
		0.000	***
	NO OF OBSERVATIONS		
	LOG-LIKELIHOOD	-9623	

NOTES: This table lists parameter estimates from the full model with search frictions. Panel A shows the parameters on the observed variables entering each of the income equations. This clarifies that the model allows for a complete set of choice-specific parameters for observables. Panel B provides estimates for the parameters describing unobserved heterogeneity. *** Significance on 1%-level; **

APPENDIX TABLE 2: MODEL MATCH - FRACTION OF OCCUPATIONAL MOBILITY ASSOCIATED WITH WAGE RISES/FALLS

		Wage Increase				Wage Decrease			
		Data	Roy	Full Model, No Frictions	Full Model, With Frictions	Data	Roy	Full Model, No Frictions	Full Model, With Frictions
Upward Mobility	Bad to Mediocre	0.470	0.455	0.777	0.602	0.349	0.466	0.151	0.339
	Bad to Good	0.632	0.372	0.774	0.640	0.223	0.560	0.175	0.287
	Mediocre to Good	0.611	0.408	0.771	0.647	0.239	0.528	0.170	0.283
Downward Mobility	Good to Mediocre	0.487	0.594	0.834	0.354	0.317	0.345	0.121	0.604
	Good to Bad	0.583	0.609	0.769	0.406	0.261	0.331	0.160	0.527
	Mediocre to Bad	0.492	0.509	0.776	0.487	0.351	0.421	0.165	0.418

NOTES: This table shows the fraction of occupational changes that are associated with discrete wage increases/decreases, as observed in the actual and simulated data. Model data are constructed from a set of 10,000 individuals for 22 years and replicates the demographic composition of the data. Observations in the simulated data that are for years past 2004 - the most recent sample year - are dropped. "Upward Mobility" is defined as occupational mobility into a better occupation, and "Downward Mobility" is defined as occupational mobility into a worse occupation. The algorithm allocating 3-digit occupations into the three groups refered to as "bad", "mediocre" and "good" is described in the text. A discrete wage change is defined as a wage change that is larger than 10 percent of the standard deviation of wages in the sample.

APPENDIX TABLE 3: MODEL MATCH - EARNINGS MOBILITY

Panel 1: 1-Year Transition Matrices

DATA (Period t+1)

Period t	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	0.798	0.144	0.030	0.017	0.011
Quintile 2	0.151	0.653	0.161	0.027	0.007
Quintile 3	0.034	0.174	0.608	0.163	0.021
Quintile 4	0.020	0.021	0.179	0.642	0.138
Quintile 5	0.013	0.006	0.017	0.145	0.820

ROY-MODEL (KEANE & WOLPIN) (Period t+1)

Period t	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1	*0.594*	*0.214*	*0.116*	*0.055*	*0.021*
Quintile 2	*0.212*	*0.308*	*0.244*	*0.154*	*0.082*
Quintile 3	*0.116*	*0.235*	*0.248*	*0.222*	*0.179*
Quintile 4	*0.054*	*0.157*	*0.217*	*0.275*	*0.297*
Quintile 5	*0.021*	*0.086*	*0.176*	*0.295*	*0.422*

FULL MODEL WITHOUT FRICTIONS

Period t					
Quintile 1	0.690	0.220	0.060	0.018	0.011
Quintile 2	0.245	0.427	0.242	0.071	0.015
Quintile 3	0.053	0.273	0.391	0.240	0.042
Quintile 4	0.006	0.074	0.269	0.453	0.198
Quintile 5	0.000	0.004	0.039	0.220	0.737

FULL MODEL WITH FRICTIONS

Period t					
Quintile 1	0.647	0.218	0.063	0.039	0.033
Quintile 2	0.238	0.475	0.238	0.045	0.004
Quintile 3	0.057	0.256	0.427	0.232	0.029
Quintile 4	0.029	0.045	0.243	0.474	0.208
Quintile 5	0.026	0.004	0.029	0.212	0.730

Panel 2: 5-Year Transition Matrices

DATA

		Period t+5				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.617*	*0.222*	*0.086*	*0.043*	*0.032*
	Quintile 2	*0.247*	*0.405*	*0.217*	*0.096*	*0.035*
Period t	Quintile 3	*0.089*	*0.251*	*0.366*	*0.218*	*0.076*
	Quintile 4	*0.045*	*0.089*	*0.252*	*0.405*	*0.207*
	Quintile 5	*0.025*	*0.028*	*0.069*	*0.230*	*0.648*

ROY-MODEL (KEANE & WOLPIN)

		Period t+5				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.567*	*0.222*	*0.125*	*0.063*	*0.023*
	Quintile 2	*0.219*	*0.243*	*0.260*	*0.174*	*0.104*
Period t	Quintile 3	*0.107*	*0.286*	*0.225*	*0.204*	*0.178*
	Quintile 4	*0.054*	*0.211*	*0.200*	*0.253*	*0.282*
	Quintile 5	*0.027*	*0.094*	*0.179*	*0.295*	*0.404*

FULL MODEL WITHOUT FRICTIONS

	Quintile 1	0.550	0.229	0.113	0.064	0.044
	Quintile 2	0.272	0.332	0.227	0.115	0.053
Period t	Quintile 3	0.113	0.274	0.310	0.223	0.081
	Quintile 4	0.027	0.134	0.276	0.363	0.200
	Quintile 5	0.002	0.017	0.078	0.251	0.651

FULL MODEL WITH FRICTIONS

	Quintile 1	0.532	0.247	0.112	0.064	0.045
	Quintile 2	0.258	0.383	0.245	0.091	0.024
Period t	Quintile 3	0.098	0.248	0.336	0.240	0.078
	Quintile 4	0.052	0.085	0.236	0.384	0.243
	Quintile 5	0.039	0.025	0.070	0.232	0.634

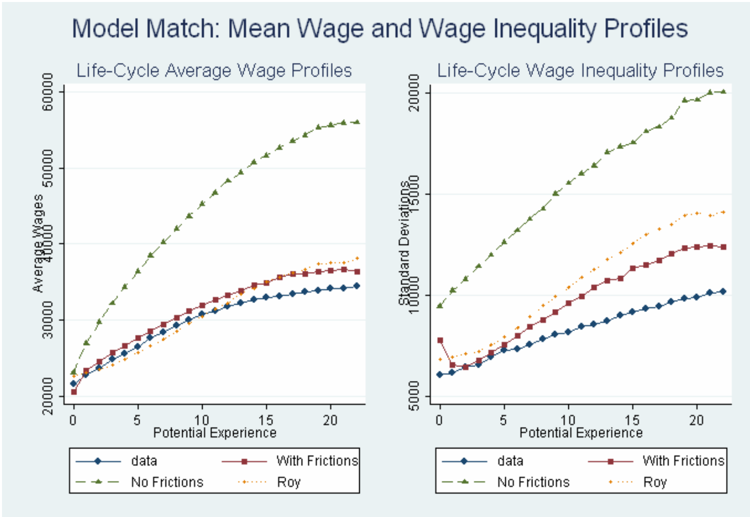
Panel 3: 10-Year Transition Matrices

		DATA					ROY-MODEL (KEANE & WOLPIN)				
		Period t+10					Period t+10				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.494*	*0.249*	*0.130*	*0.074*	*0.053*	*0.548*	*0.219*	*0.136*	*0.071*	*0.027*
	Quintile 2	*0.255*	*0.312*	*0.227*	*0.141*	*0.064*	*0.189*	*0.294*	*0.237*	*0.167*	*0.114*
Period t	Quintile 3	*0.131*	*0.239*	*0.282*	*0.228*	*0.120*	*0.108*	*0.217*	*0.232*	*0.226*	*0.218*
	Quintile 4	*0.073*	*0.126*	*0.237*	*0.314*	*0.249*	*0.061*	*0.150*	*0.208*	*0.268*	*0.313*
	Quintile 5	*0.035*	*0.051*	*0.111*	*0.253*	*0.551*	*0.027*	*0.101*	*0.195*	*0.299*	*0.378*

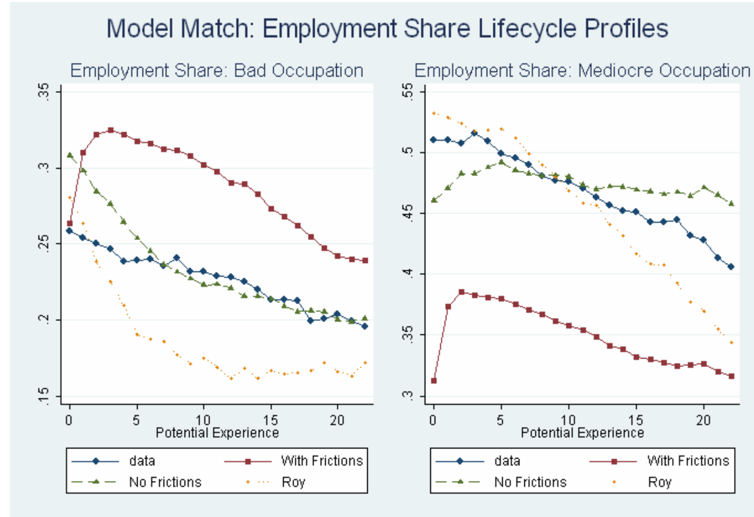
		FULL MODEL WITHOUT FRICTIONS					FULL MODEL WITH FRICTIONS				
	Quintile 1	0.420	0.227	0.156	0.112	0.086	0.442	0.255	0.146	0.093	0.064
	Quintile 2	0.271	0.275	0.209	0.147	0.098	0.263	0.315	0.237	0.130	0.054
Period t	Quintile 3	0.161	0.252	0.260	0.207	0.121	0.133	0.234	0.275	0.232	0.126
	Quintile 4	0.066	0.178	0.260	0.298	0.198	0.066	0.122	0.227	0.321	0.264
	Quintile 5	0.010	0.045	0.122	0.268	0.555	0.047	0.048	0.115	0.248	0.542

NOTES: This table shows the probability that a worker whose income falls into the q-th quintile of experience-specific income distributions receives income in the p-th quintile 1 year, 5 years and 10 years later, as observed in the actual and simulated data. Model data are constructed from a set of 10,000 individuals for 22 years and replicates the demographic composition of the data. Observations in the simulated data that are for years past 2004 - the most recent sample year - are dropped.. Wages are corrected from cohort effects.

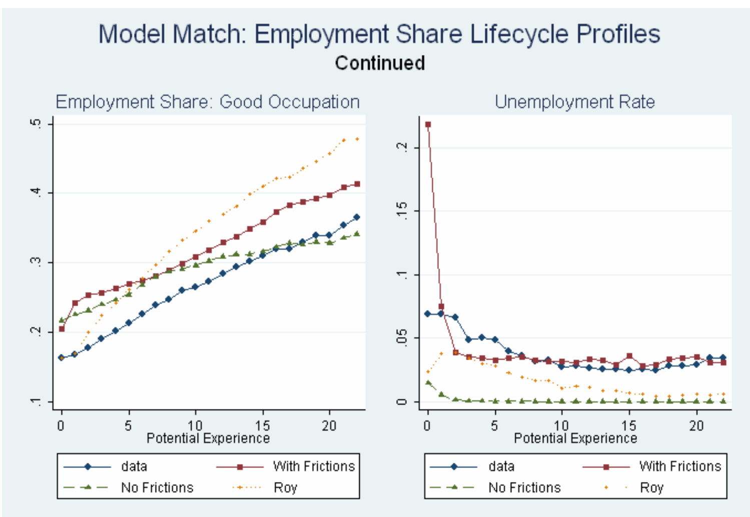
APPENDIX FIGURE 1



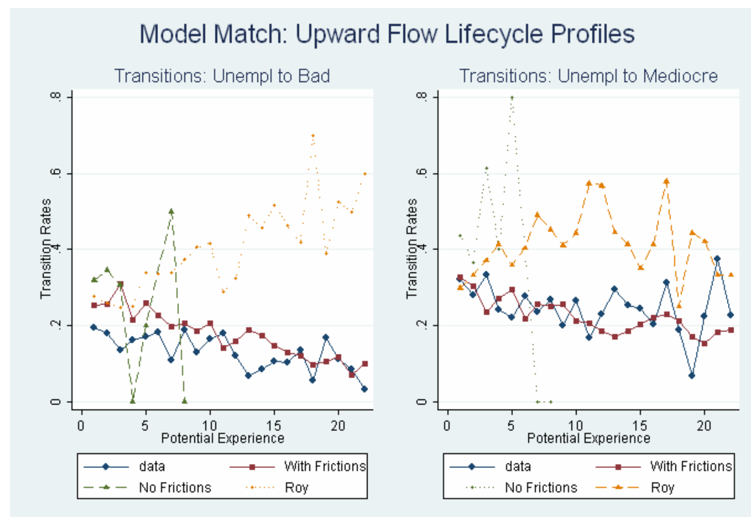
APPENDIX FIGURE 2



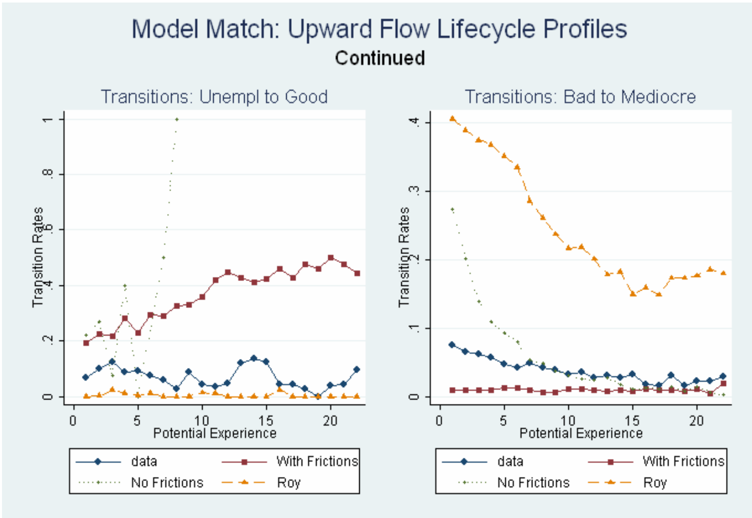
APPENDIX FIGURE 3



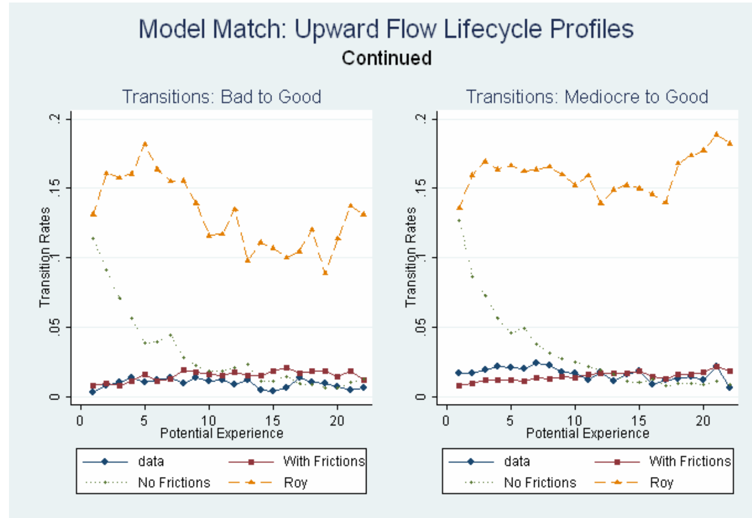
APPENDIX FIGURE 4



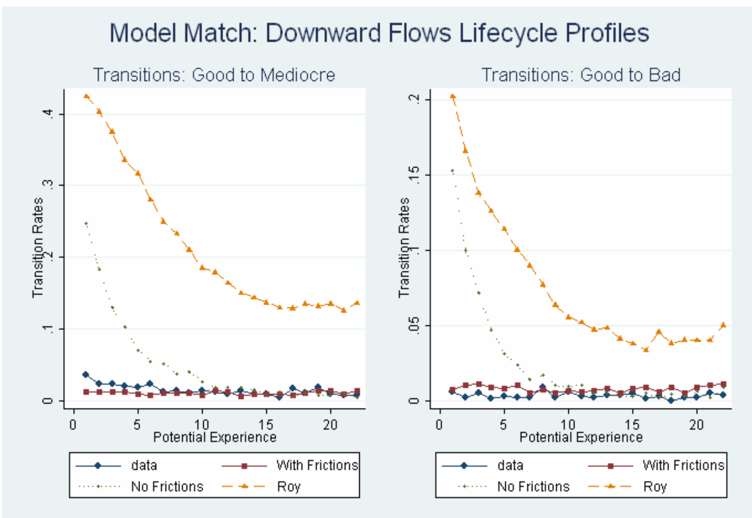
APPENDIX FIGURE 5



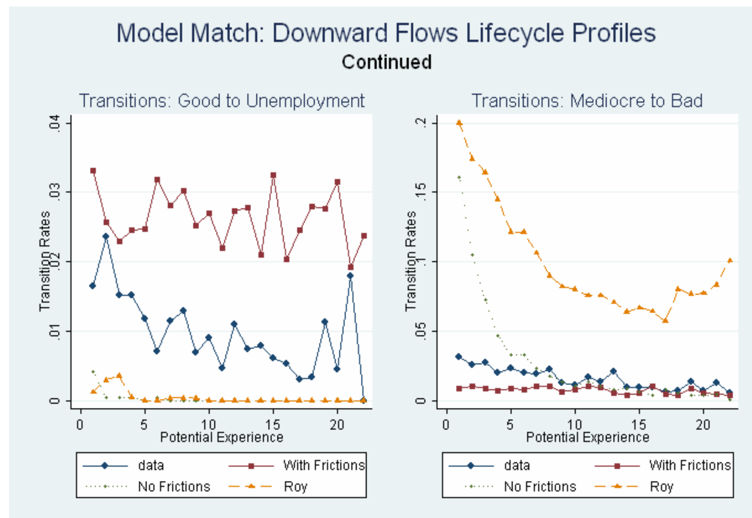
APPENDIX FIGURE 6



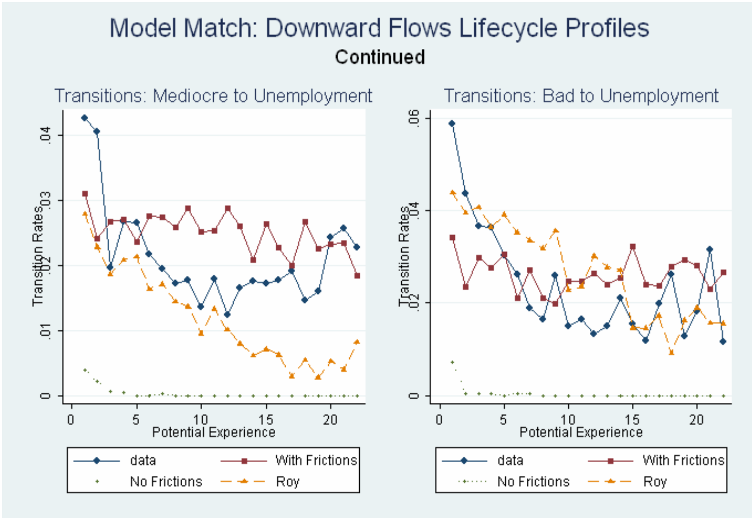
APPENDIX FIGURE 7



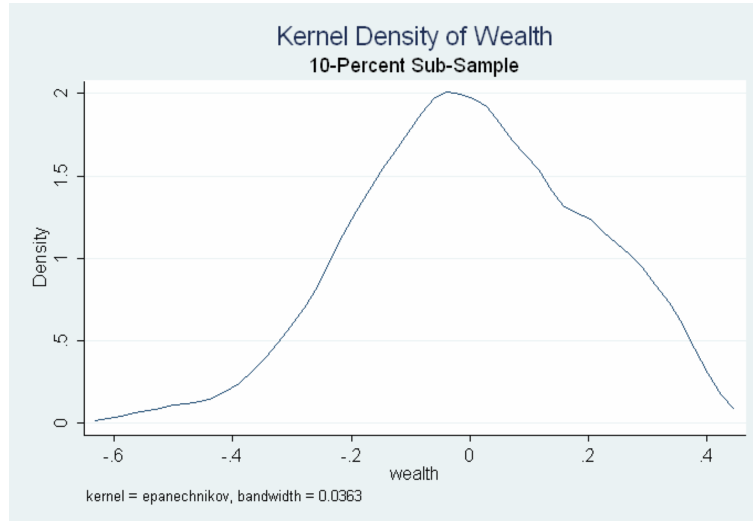
APPENDIX FIGURE 8



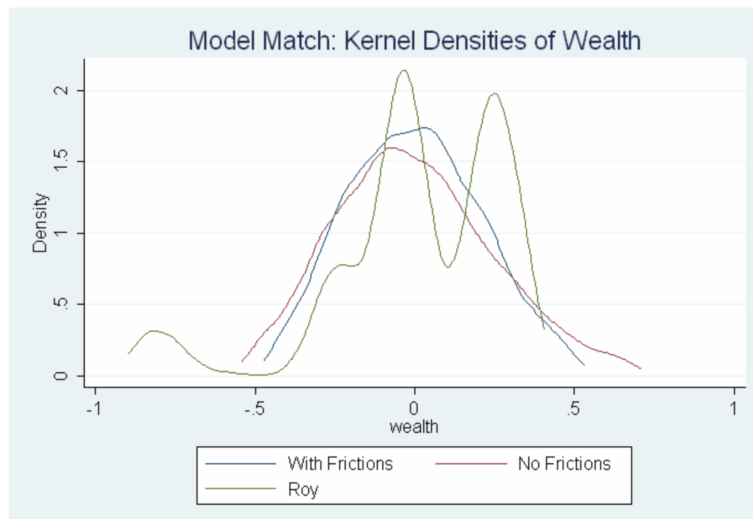
APPENDIX FIGURE 9



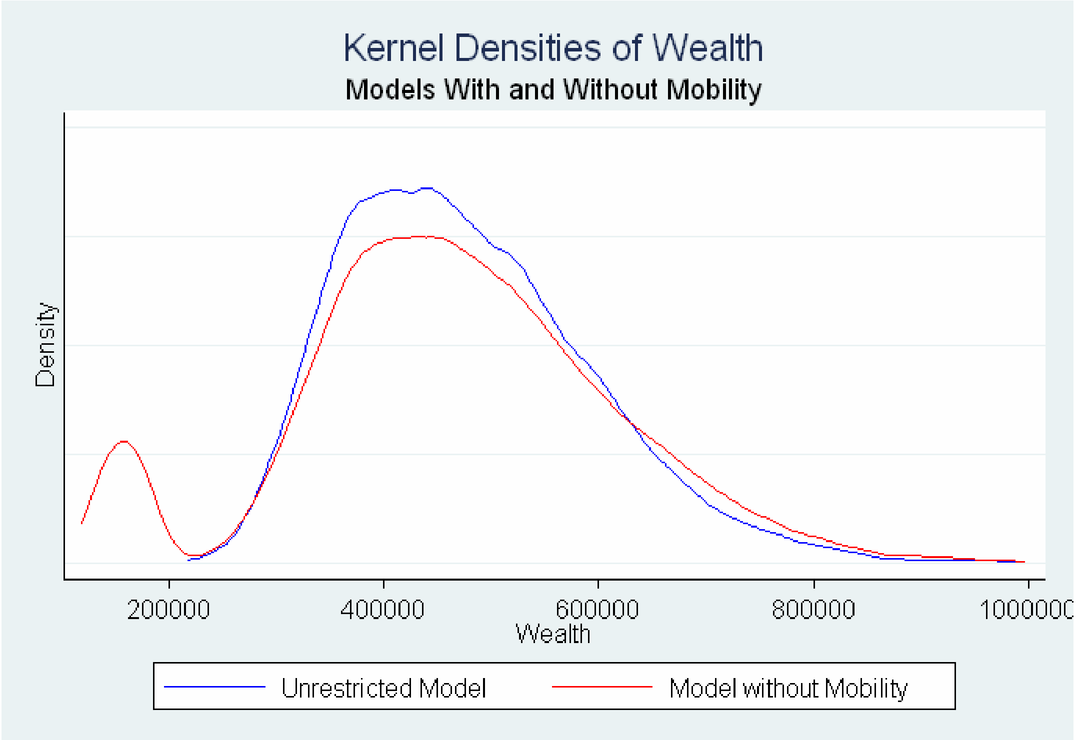
APPENDIX FIGURE 10 (See Figure 12)



APPENDIX FIGURE 11



APPENDIX FIGURE 12



APPENDIX FIGURE 13