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

Florian Hoffmann
Department of Economics
University of British Columbia

PRELIMINARY VERSION
August 2010

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, 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. It remains highly controversial if this career heterogeneity is predominantly driven by unobserved skills acquired prior to labor market entry, such as parental investments or innate abilities, or by decisions made and shocks accumulated over the working life, such as "luck" to job search outcomes, business cycle fluctuations or structural changes. 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 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 (1997) 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. 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 they are hit by entirely different sequences of exogenous skill shocks.[3] Furthermore, mobility itself generates permanent changes in wages and earnings that are not necessarily related to comparative advantages established prior to labor market entry, but rather to "luck" of finding a particularly good match, to geographic differences in occupational average wages, or to un-modeled differences in tasks within occupational classes. Consequently, when these sources of dynamic unobserved heterogeneity are not controlled for in the structural model, they will be erroneously subsumed

[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).

[3] Any permanent change in the environment that is not endogenous or observable and translates into wage changes will be included in these "permanent shock". One may think of health shocks, firm closures, changes in the family structure, and promotions or demotions within occupational classes.

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 occupational match heterogeneity and exogenous permanent shocks to skills. Match effects remain constant for the duration of a job.[4] This generates a jump process of residual wages, with jumps occurring 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.

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 Simulated Maximum 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 significantly 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 28 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"

[4] 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).

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 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.

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 increasing earnings. These conclusions only apply for ability differences within educational groups and do 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.[5] Sullivan (2010) enriches the Keane-Wolpin framework with inter-firm mobility and firm-specific human capital. To keep the model computationally tractable he needs to impose a number of strict assumptions. Most importantly, both, occupational and firm-specific human capital, take on only three values that are randomly updated. Thus, the endogenous evolution of human capital endowments is drastically simplified. Two recent applications, Adda, Dustmann, Meghir and

[5]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).

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 and a selection equation for firm mobility.[6]

The rich specification of the institutional environment and of unobserved heterogeneity 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.[7] 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.[8]

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. Computational tractability restricts the complexity of these models considerably. On the one hand, a Dynamic Programming Problem needs to be solved at each value of the parameter vector considered in the optimization routine. On the other hand, the likelihood usually does not admit a closed-form and thus needs to be simulated. 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 Dynamic Discrete Choice framework. Given that the state-space becomes intractably large, the resulting model cannot be estimated by conventional algorithms that solve the Dynamic Programming Problem in the inner routine and that maximizes the likelihood in the outer routine. In fact, unless one has access to large computer clusters, the Dynamic Programming Problem cannot be solved on a reasonable number of discretization points even once.[9] Instead I utilize an estimation algorithm that has been proposed by Geweke and Keane (2000) and that relies on parametric approximations of expected choice-specific values.

[6]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 employment is associated with one wage equation only. Firm mobility merely changes the random intercept, but not the human-capital endowment. 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.

[7]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).

[8]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, unserved heterogeneity is still very restrictive.

[9]There is a large and quickly growing literature in empirical IO that develops methods that circumvent the "curse of dimensionality" in Dynamic Discrete Choice models. All of these methods rely on the linearity of payoff functions and a particular error structure, both of which do not apply to my framework. See e.g. Hotz and Miller (1993), Hotz, Miller, Sanders and Smith (1994), Aguirregabiria and Mira (2002, 2007), Imai, Jain and Ching (2006), Norets (2009), and Arcidiacono and Miller (2010). Su and Judd (2008) propose an alternative and potentially fast approach based on ideas from constrained optimization. In this approach the Bellman-operator still needs to be evaluated repeatedly.

With data on earnings and choices, the parameters of this approximation function can be estimated without solving the Dynamic Programming Problem even once. At the same time, the model is still fully consistent with agents to be forward looking when making optimal choices about human-capital investments and occupational choices.[10]

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 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. 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

[10]Parametric approximations of expectations have also been used frequently in the computation of dynamic macro models. With continuous control variables, one approximates the expectations in stochastic Euler-equations instead of expected values. While the approximation parameters are identified from observed choices in Dynamic Discrete Choice models, they are identified from equilibrium conditions in DSGE-applications. See e.g. den Haan and Marcet (1990), Marcet and Marshall (1992, 1994), Marcet and Lorenzoni (1999), Christiano and Fisher (2000), Duffy and McNelis (2001) and Heer and Maussner (2008). Similar ideas, though with a bounded rationality interpretation, are applied in Krussell and Smith (1998) and Lee and Wolpin (2006).

[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 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).

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 subsequent section I lay out the model structure and discuss its main features. Estimation, identification and the Geweke-Keane algorithm are 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 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.

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

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. Modeling 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.

[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.

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' 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 career trajectories, often seen as reflecting "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 "blue-collar", "pink-collar" and "white-collar". 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 labeled as "blue-collar" and the one third of occupations with the highest proportions of highly educated are labeled as "white-collar". The rest are labeled as "pink-collar". Mechanically dividing

[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.

the support of the distribution of occupation-level proportions of highly educated workers serves to keep sample sizes for each occupation group relatively large.

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.[18] The former group is entirely covered by blue-collar occupations, while the latter is entirely contained within the white-collar 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 blue-collar group, but the majority falls into the pink-collar 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 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 carried over to the sub-sample, but with considerably more noise around the trend line.

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 3 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 3). 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 4).

2. There is a large re-allocation of workers from blue-collar and pink-collar occupations to white-collar occupations over the life-cycle as reflected by a significant decrease of the employment shares in the

[18] This relationship is statistically significant.

two former occupational classes and a strong increase in the latter (figure 5).

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 white-collar occupations compared to the other two occupation groups (figures 6 and 7).

4. Upward mobility is most pronounced for transitions from blue- to pink-collar occupations, followed by transitions from pink- to white-collar occupations. There is only very little mobility from blue- to white-collar occupations (figure 8).

5. In the other direction, transition rates from white- to pink–collar occupations and from pink- to blue-collar occupations are frequent, but decreasing over the life-cycle. Mobility from white- to pink-collar occupations is almost non-existent (figure 9).

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 discounted sum of earnings net of cohort effects, over the first 23 years of individual careers is right-skewed (figure 10).

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 (table 3).

   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, white-collar 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 blue-collar or from blue-collar directly to white-collar. However, a large fraction of individuals, having acquired a vocational training degree by construction of the sample, start directly in the pink-collar occupation at the time of labor market entry.

3. The Model

3.1. General Description

The model merges the dynamic discrete choice framework of occupational mobility and endogenous human capital accumulation by Keane and Wolpin (1997) with a flexible variance components model of residual wages. Individuals maximize life-time utility by choosing among four alternatives in each period. To abstract from the savings decision I follow the majority of studies based on discrete choice models and assume that workers are risk-neutral. The choices, indexed by $j \in \{u, bc, pc, wc\}$, are unemployment (u) and working in a blue-collar (bc), pink-collar (pc) or white-collar (wc) occupation. Individuals, whether employed or not, receive wage offers from each occupation group in each period. Wages are determined by skill prices and a skill index that is composed of human-capital endowments, unobserved person-specific comparative advantages, a random-walk for general skills, transitory occupation-specific shocks, and a match-specific component. Human capital includes a general and an occupation specific part, both evolving endogenously over time. Each alternative is associated with a time-constant non-monetary utility component B^j. To keep the model as tightly parameterized as possible I assume these parameters to be non-stochastic. Identification requires to normalize one of these parameters, and I choose to set $B^{wc} = 0$.

When moving into an occupation, a worker needs to incur a fixed cost of mobility. [19] Since labor market entrants, unemployed workers and employed workers potentially face different frictions I assume mobility costs to be c^u for the first group, c^{enter} for the second group, and c^{empl} for the third group. Matches are exogenously broken up at a rate δ, forcing the worker into unemployment.

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.

3.2. Occupation Specific Wages

The potential wage of individual i 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 $j \in \{bc, pc, wc\}$. Log-wages are therefore equal to $p_t^j + h_{it}^j$, 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:

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

[19] An earlier version of this paper replaced the fixed-costs to mobility with explicit search frictions, i.e. Poisson-rates at which workers draw offers. In the presence of search frictions, if a worker does not receive offers from all occupations, or if his only choice is between staying in the current occupation or becoming unemployed, a worker needs to solve a constrained optimization problem. Therefore, each additional friction introduces an additional value function, thus increasing the computation costs in the estimation routine.

$$p_t^j = \alpha_5^j * t + \alpha_6^j * (t)^2 + \alpha_7^j * U_t + \alpha_8^j * (U_t)^2 \quad (3.2)$$

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

$$x_{it+1} = x_{i,t} + 1(j_i^*(t) \in \{bc, pc, wc\}); \quad x_{i0} = 0 \quad (3.3)$$

$$ten_{it+1}^j = ten_i^j + 1(j_i^*(t) = j); \quad ten_{i0}^j = 0. \quad (3.4)$$

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. 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 essentially describe the Keane and Wolpin (1997) model. I extent their framework 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, capturing general equilibrium effects operating through demand side shocks.[20] In particular, the model allows occupation specific skill prices to trend and to react

[20] 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.

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.[21] Both components introduce exogenous variation into the model and the choice rules of optimizing agents.

3.3. Unemployment Benefits

Unemployment is an alternative 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 drop to the ALH level. The maximum length of AG-collection depends on the age of an individual. For the cohort- and age-groups present in my sample, this limit has never exceeded one year during the sample period. Both, AG and ALH also depend on past wages earned on the job.[22] There have been several reforms of the unemployment system during the sample period that either affected levels of AG and ALH or the income replacement rate.

In theory, the parameters governing the unemployment benefits an individual is entitled to are set by institutions and can be directly inferred from the code of labor law. In practice, the rules and exceptions are too complicated to be parameterized in a transparent way, and the data are potentially affected by measurement error. Thus, I choose a parametric specification that incorporates the crucial characteristics of the unemployment insurance system and its evolution over time and estimate the parameters from the data. 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 $1(dur_{it} \geq 1)$ a dummy equal to one if unemployment duration is at least one years, in which case AG drops to the ALH level. The following equation matches observed unemployment benefits $\widetilde{w}^u_{it}$ extremely well $\left(R^2 > .99\right)$:

$$\begin{aligned}\widetilde{w}^u_{it} \quad = & \left[\begin{array}{c}\alpha^u_{0,1983} * 1(t \leq 1983) + \alpha^u_{0,84\_93} * 1(1984 \leq t \leq 1993) \\ +\alpha^u_{0,94\_04} * 1(1994 \leq t \leq 2004)\end{array}\right] \\ & +w_{i,-1} * \left[\begin{array}{c}\alpha^u_{1,1983} * 1(t \leq 1983) + \alpha^u_{1,84\_93} * 1(1984 \leq t \leq 1993) \\ +\alpha^u_{1,94\_04} * 1(1994 \leq t \leq 2004)\end{array}\right]\end{aligned}$$

[21] 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, Jacobeb-binghaus and Ludsteck (2009) replicate the study using the IABS and reach at similar conclusions.

[22] 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).

$$
+1(dur_{it} \geq 1) * w_{i,-1} * \left[\begin{array}{c} \alpha^u_{2,1983} * 1(t \leq 1983) + \alpha^u_{2,84\_93} * 1(1984 \leq t \leq 1993) \\ +\alpha^u_{2,94\_04} * 1(1994 \leq t \leq 2004) \end{array} \right]
$$

$$
+\omega_{it}. \tag{3.5}
$$

where ω_{it} is a classical measurement error. Unemployment benefits are determined by a base amount α^u_0, a replacement rate α^u_1 under AG, and an adjustment of the replacement rate of α^u_2 when AG drops to ALH. All parameters vary freely across three major periods which are marked by reforms to the unemployment insurance system. Since individuals do not act on ω_{it}, this equation can be estimated by simple OLS using unemployment benefit data for a sample of all unemployment benefit recipients below an age of 45, the highest age contained in my working sample.[23] Only the deterministic part will enter the discrete choice model to be estimated below. It is important to highlight that unemployment duration evolves endogenously according to the law of motion

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

which will be taken into account in the individual's decision problem below.

I introduce a stochastic component to unemployment benefits by introducing a non-monetary utility component ε^u_{it} that is, in contrast to B^u, random in the population.

3.4. Stochastic Structure

The model features several variance components of unobserved heterogeneity. Individuals are endowed with occupation-specific skill levels $\alpha^j_{0,i}$ that are hit by transitory shocks ε^j_{it} 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 are discretely distributed. In particular, as clarified by equation (3.7), I assume that there are K types of individuals in the population, each endowed with a vector $\left(\alpha^{bc}_{0,k}, \alpha^{pc}_{0,k}, \alpha^{wc}_{0,k}\right)_{k=1,\ldots K}$ 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^j_{it} 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 initialize the stochastic process at zero. The dynamics of match heterogeneity is described in equation (3.10), forcing match effects to be constant for the

[23] This approach is similar to calibrating the parameters to moments describing the unemployment benefit system.

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:

$$
\begin{aligned}
Prob\left(\alpha_{0,i}^{bc}=\alpha_{0,k}^{bc},\alpha_{0,i}^{pc}=\alpha_{0,k}^{pc},\alpha_{0,i}^{wc}=\alpha_{0,k}^{wc}\right) &= \pi^{k},\sum_{k}^{K}\pi^{k}=1, K=4 && (3.7)\\
\varepsilon_{it}^{j} &\sim N(0,\sigma_{\varepsilon,j}^{2}) && (3.8)\\
u_{it} &= u_{it-1}+\xi_{it},\ \left\{\begin{array}{c} u_{i0}=0 \\ \xi_{it}\sim N(0,\sigma_{\xi}^{2})\end{array}\right\}, && (3.9)\\
\mu_{it}^{j\in\{bc,pc,wc\}} &= \left\{\begin{array}{c} \mu_{i,t-1}^{j}\ if \quad j(t)=j_{i}^{*}(t-1) \\ \nu_{it}^{j}\sim N(0,\sigma_{\nu}^{2})\text{ otherwise}\end{array}\right\} && (3.10)
\end{aligned}
$$

3.5. Bellman Equations

At the beginning of each period, individuals have to choose among the four alternatives $j \in \{u, bc, pc, wc\}$. They observe their potential alternative-specific log-wages $w_{it}^{j\in\{bc,pc,wc\}} = p_t^j + h_{it}^j$, with p_t^j and h_{it}^j given by equations (3.2) and (3.1) respectively, and unemployment benefits w_{it}^u, given by (3.5). Both, employers and employees are perfectly informed about the current state. Given wages, unemployment benefits, mobility costs, non-monetary benefits, and the current individual-specific states, employees choose the alternative with the highest expected payoff. To concisely summarize the decision problem, let $S_{it} = \left(X_{it}, 1(j=j^*(t)), k, \mu_{it}^j, u_{it}, \varepsilon_{it}^j\right)_{j\in\{u,bc,pc,wc\}}$ be the period-t state vector for individual i, composed of the observable variables X_{it}, a list of indicator variables equal to one if alternative j is the optimal choice for individual i in period t, 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. The dummy variables $1(j = j^*(t))$ need to be included in the state-space as well as they influence the expected fixed-costs to mobility in the next period.

With discount factor β, the Dynamic Programming Problem is thus described by

$$
\begin{aligned}
V_{it}^{j(t)\neq u}(S_{it}) &= \exp\left\{w_{it}^{j}\right\}+B^{j}+1(age_{it}=23)*c^{entry}+1(j(t)\neq j^{*}(t-1)=u)*c^{u}\\
&\quad +1(j(t)\neq j^{*}(t-1)\neq u)*c^{emp}
\end{aligned}
$$

$$+\beta * (1-\delta) * E\left[V_{it+1}(S_{it+1}) | S_{it}, j^*(t) = j\right]$$

$$+\beta * \delta * E\left[V^u_{it+1}(S_{it+1}) | S_{it}, j^*(t) = j\right] \quad (3.11)$$

$$V^u_{it}(S_{it}) = \exp\{w^u_{it}\} + B^u + \beta * E\left[V_{it+1}(S_{it+1}) | S_{it}, j^*(t) = u\right] \quad (3.12)$$

$$V_{it}(S_{it}) = \max\left\{V_{it}^{j(t)\in\{u,bc,pc,wc\}}(S_{it})\right\} \quad (3.13)$$

$$S_{it+1} = \Lambda(S_{it}, j^*(t)) \quad (3.14)$$

Equation (3.11) is the value function for employment. In the current period, individuals receive potential wages $\exp\left\{w^j_{it}\right\}$ and non-monetary rewards B^j and need to pay fixed costs of mobility that depend on the previous choice. In the next period, with probability $(1-\delta)$ they can choose freely among all alternatives, and with probability δ they are forced into unemployment. Equation (3.12) is the value function for unemployment, an alternative that can be chosen in any period. Unemployed individuals collects unemployment benefits w^u_{it} and receive wage offers from all occupation groups in the subsequent period. Equation (3.13) is the Bellman-equation for the unconstrained optimization problem. The final equation is the updating rule of the state space. General experience x_{it}, occupation specific tenure ten^j_{it}, unemployment duration dur_{it}, the random walk u_{it} and the match effects μ^j_{it} are updated according to the dynamic equations (3.3), (3.4), (3.6), (3.9) and (3.10).

3.6. Further Discussion of Model Features

Human capital accumulation directly influences mobility decisions. General human capital, having occupation-specific returns, 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.

The empirical model described above introduces two crucial features to the Keane-Wolpin framework: First, 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. On the one hand, 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. On the other hand, 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 create a downward bias of the estimated dispersion of permanent shocks. Only very few studies, most notably Altonji et al. (2009) and Low et. al (2009), correct for this bias, albeit in reduced form frameworks. Finally, 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: Match heterogeneity induces a jump-process in residual earnings at the time of occupational mobility, consistent with reduced form evidence of a systematic relationship between earnings and mobility. When estimating unit-roots models without controlling for mobility, residual wage fluctuations that induce behavioural adjustments are mistakenly interpreted as exogenous and neutral shocks, therefore leading to an upward bias of random walk shocks.

A second difference to the Keane-Wolpin model is the presence of within-occupation match heterogeneity. This model feature is what introduces a systematic and endogenous relationship between mobility and residual earnings dynamics. Consequently, individuals with the same pre-labor market skills and thus identical entry wages and long-run occupational sorting – the two data features essentially identifying type heterogeneity – can have different discrete earnings adjustments at the time of mobility and thus introduces within-type inequality. Occupational match quality as considered in this paper subsumes any residual wage component that is permanent for the duration of a match. As examples of factors influencing the amount of match heterogeneity one may think of geographic differences in occupational average payments or of earnings differences across 3-digit occupations that are summarized in the three broad occupational categories. Another interpretation is that match-heterogeneity absorbs the average employer-match-quality for all firms that are visited while staying in a particular occupation, while the random-walk component absorbs the residual wage fluctuations around the "average firm-match" generated by inter-firm mobility within an occupational group.

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 $X_{it} = \left(t, U_{it}, x_{it}, ten_{it}^{j\in\{bc,pc,wc\}}, dur_{it}, w_{i,-1}, age_{it}\right)$ is the vector of predetermined variables. To spare on notation, I also define $\mu_{it} = \left(\mu_{it}^j\right)_{j\in\{bc,pc,wc\}}$ to be the full vector of match effects, $\widetilde{S}_{it}$ 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 exogenous job breakups, non-monetary benefits and mobility costs. These model features are straightforward to introduce.

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.

4.1. The Likelihood

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

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

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:

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

where ϕ is the pdf and Φ is the cdf of the Standard Normal Distribution. The first term is the conditional wage density, collapsing to one mass point if the choice is unemployment, 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.[24] In particular, given the discreteness of α_i we have

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

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. I follow the simulation algorithm outlined in Sullivan (2010).

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

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

Standard errors are computed from the inverse Hessian.

4.2. The Geweke-Keane Estimator and its Interpretations

To estimate their Dynamic Discrete Choice Model, Keane and Wolpin utilize a double-nested algorithm that solves the Dynamic Programming Problem at a number of state-points, then interpolates the value function between state-points using a regression with the values as dependent and state-points as independent variable, and finally constructs the likelihood. This step needs to be performed at each value of the structural parameters that is chosen by the maximization routine, including perturbations performed to obtain the numerical gradient. In practice, the slow convergence properties of the estimator implies that the Dynamic Programming, interpolation, and computation of the likelihood need to be carried out thousand of times. With the introduction of dynamically evolving error components, i.e. the 3-dimensional match components and the random walk, such an estimation routine is numerically infeasible on all but the strongest computers.

[24] 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).

This is an expression of the well-known Curse-of-Dimensionality. Alternative algorithms that are used in IO to solve this problem, such as Hotz and Miller's (1993) CCP-estimator, cannot be used in applications with non-linear payoffs and the type of unobserved heterogeneity considered in my paper. Instead I rely on an estimator proposed by Geweke and Keane (2000) in which the Dynamic Programming Problem does not need to be solved even once. The estimator is based on ideas from functional approximation theorems that show that any continuous function can be approximated by a sufficiently long polynomial. One therefore parameterizes the expectation functions $E\left[V_{it+1}\left(S_{it+1}\right)|S_{it}, j(t)\right]$ and $E\left[V^u_{it+1}\left(S_{it+1}\right)|S_{it}, j(t)\right]$ as polynomials of the state-variables and infers the parameters from the data. The approximations are fully consistent with the evolution of the state-space under each choice.

With both, payoff and choice data, one can identify all parameters that are preserved under taking differences of the polynomials. Consider an example in which an individual picks the alternative with the lowest wage. Assuming forward-looking behavior this implies that the parameters of the approximation functions need to adjust such that the approximated value from choosing this alternative is sufficiently large to induce the observed behavior. In other words, the parameters of the approximation function are identified from choices that are hard to explain with differences in one-period payoffs.

Before providing a simple example I discuss three possible interpretations of the parameters. First, one may think of it as an approximation to the full Dynamic Discrete Choice Estimator that involves solutions of the Dynamic Programming Problem. In fact, the double-nested estimator used in Keane and Wolpin (1997) also computes a polynomial approximation function for interpolation. However, while in their estimation routine the parameters are computed from solving the Bellman-equation on a tractable number of state-points and then using the values as "independent variable" in regressions, the parameters of the approximation function of the Geweke-Keane estimator are identified from observed choices, net of differences in alternative-specific wages. Clearly, since this algorithm never computes "data" from the Dynamic Programming Problem, less parameters of the approximations can be identified. I discuss this point further below. It is important to note that with fully rational individuals any polynomial of a finite degree will introduce finite-sample bias to the estimate. Only extensive Monte-Carlo analysis can provide evidence about the size of this bias. Monte-Carlo analysis in Geweke and Keane (2000) and my own work suggests that with a sufficiently high degree this bias is negligible.[25]

A second natural interpretation of the estimator is in terms of bounded rationality. With this interpretation, economic agents, although still forward looking, use some heuristic to calculate the expected values of choosing an alternative instead of solving the full Dynamic Programming Problem. The Geweke-Keane estimator postulates that the individuals use simple functions to compute very complicated expectations.[26]

[25] One approach is to simulate Monte-Carlo data with fully rational individuals, to estimate the parameters using the Geweke-Keane estimator and then to compare life-cycle earnings that are generated with the policy functions from the Dynamic Programming problem with those generated from policies when using polynomial approximations. Geweke-Keane show these differences to be negligible.

[26] Parameterizing expectations has a significantly longer history in the computational DSGE-literature than in Dynamic Discrete Choice Modeling. See Heer and Maussner (2008), Chapter 5 for a discussion. This reference also provides a formal link between the interpretation as an approximation to the full Dynamic Programming model and the interpretation as a model with bounded rationality.

Lastly, one can also interpret the estimator as rationalizing a flexible parametric specification of conditional choice probabilities. Since there is little guidance as to the right degree of the polynomials to be used in the approximation, one adds higher degrees until the increase in the likelihood function is insignificant. Given that future states enter these choice probabilities through the laws of motions generated by each choice while they do not enter the one-period payoff functions, the Geweke-Keane estimator provides exclusion restrictions. Intuitively, differences in next-period states are a result of choices and thus enter the conditional choice probabilities while they do not enter current wages.

For the empirical application in this paper, an approximation function with 4th-order polynomials and 1st-order interactions of the state variables produces very good results.[27] All parameters are allowed to vary freely by type. Since individuals have a finite horizon, interactions with age are particularly important as they capture the evolution of values as an individual ages. As discussed below, only a subset of parameters can be identified.

4.3. Example

To demonstrate the Geweke-Keane estimator consider an example with only two alternatives to chose from, occupations 1 and 2, no exogenous job breakups, a non-monetary utility benefit of working in occupation 1, B^1, a cost of mobility, c, that is the same for labor market entrants and employed individuals, and the following wage equations for occupation j:[28]

$$w_{it}^j = \alpha_0^j + \alpha_1^j * ten_{it}^1 + \alpha_2^j * \left(ten_{it}^1\right)^2 + \alpha_3^j * ten_{it}^2 + \alpha_4^j * \left(ten_{it}^2\right)^2 + \varepsilon_{it}^j; \quad \varepsilon_{it}^j \tilde{} N(0, \sigma_j^2). \tag{4.5}$$

Given that there is neither unemployment nor any source of non-stationarity we can assume that t denotes potential experience rather than calendar year. Hence, $ten_{it}^1 + ten_{it}^2 = t$, and conditional on t only one tenure variable enters as a state into the Bellman equation. The remaining state-variables are an indicator function which is equal to one if $j^*(t) = 1$, $occ1_{-1}$, and the two occupation-specific unobserved shocks. Since the latter do neither directly influence the laws of motion nor contain information about the distribution of future shocks they can be excluded from the set of state variables. The Bellman-equations are given by

$$\begin{aligned} V_1(ten_{it}^1, t_i, occ1_{-1}) &= \exp(w_{it}^1) + B^1 + (1 - occ1_{-1}) * c + \beta * E\left[V\left(ten_{it}^1 + 1, t_i + 1, 1\right)\right] \\ V_2(ten_{it}^1, t_i, occ1_{-1}) &= \exp(w_{it}^2) + occ1_{-1}, *c + \beta * E\left[V\left(ten_{it}^1, t_i + 1, 0\right)\right] \\ V(ten_{it}^1, t_i, occ1_{-1}) &= \max\left\{V_1(ten_{it}^1, t_i, occ1_{-1}), V_2(ten_{it}^1, t_i, occ1_{-1})\right\}. \end{aligned} \tag{4.6}$$

[27] In total, this approximation function introduces 110 additional parameters.
[28] The example is taken from Geweke and Keane, with some modifications.

The Geweke-Keane estimator approximates $E\left[V\left(ten^1_{it}+1,t_i+1,1\right)\right]$ with a polynomial. Choosing the approximation function to be a second order polynomial with first-order interactions we obtain

$$
\begin{aligned}
& E\left[V\left(ten^1_{it}+1,t_i+1,1\right)\right]-E\left[V\left(ten^1_{it},t_i+1,0\right)\right] \\
\approx & \left[\begin{array}{c}\pi_0+\pi_1*\left(ten^1_{it}+1\right)+\pi_2*\left(ten^1_{it}+1\right)^2+\pi_3*(t_i+1)+\pi_4*(t_i+1)^2 \\ +\pi_5+\pi_6*\left(ten^1_{it}+1\right)*(t_i+1)+\pi_7*\left(ten^1_{it}+1\right)+\pi_8*(t_i+1)\end{array}\right]- \\
& \left[\begin{array}{c}\pi_0+\pi_1*ten^1_{it}+\pi_2*\left(ten^1_{it}\right)^2+\pi_3*(t_i+1)+\pi_4*(t_i+1)^2 \\ +\pi_6*ten^1_{it}*(t_i+1)+\pi_7*ten^1_{it}+\pi_8*(t_i+1)\end{array}\right] \\
= & \pi_1+\pi_2*\left(2*ten^1_{it}+1\right)+\pi_5+\pi_6*(t_i+1), \qquad (4.7)
\end{aligned}
$$

where the $\pi's$ are the parameters of the approximation function. In particular, π_5 is the parameter on $occ1_{-1}$, and π_7 and π_8 are the parameters on its interactions with other state-variables. The choice probability of choosing occupation 1, conditional on the observed wage and the state-variables, is given by

$$
P\left(\begin{array}{c}\exp(w^1_{it})+B^1+(1-2*occ1_{-1})*c \\ +\beta*\left(\pi_1+\pi_2*\left(2*ten^1_{it}+1\right)+\pi_5+\pi_6*(t_i+1)\right)>\exp(\widehat{w}^2_{it})*\exp(\varepsilon^2_{it})\end{array}\right). \qquad (4.8)
$$

This example clarifies that only parameters on state-variables and interaction terms that are affected by choices through the laws of motions can be identified. Furthermore, the fixed costs of mobility are identified even though the choices made today enter as a state variable because of the existence of such costs. Further identification results are discussed in the next session.

4.4. Identification

The model is kept tightly parameterized, and identification is transparent. Excluding the parameters of the unemployment benefits equation that is estimated separately by OLS there are 49 parameters, 21 parameters of which describe the observable part of log-income and are essentially selection-corrected regression estimates, and 15 parameters of which determine the non-parametric distribution of comparative advantage. Only 2 parameters capture the processes of the random walk and the match heterogeneity of the unobservable part, 4 parameters are variances of the transitory shocks, one parameter is the job-breakup rate, 3 parameters are fixed-costs to mobility, and 3 parameters are non-monetary benefits.

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

$$\begin{aligned} w_{it}^{j^*(t)} - w_{it-1}^{j^*(t-1)} &= X_{it}^{j^*(t)}\beta^{j^*(t)} - X_{it-1}^{j^*(t-1)}\beta^{j^*(t-1)} \\ &+\xi_{it} + 1\left[j_i^*(t) \neq j_i^*(t-1)\right] * \nu_{it}^{j} + \varepsilon_{it}^{j^*(t)} - \varepsilon_{it-1}^{j^*(t-1)} \end{aligned} \quad (4.9)$$

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 $\left(\alpha_{0,k}^{bc}, \alpha_{0,k}^{pc}, \alpha_{0,k}^{wc}\right)$ 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. Panel data enable a researcher to estimate occupation-specific intercepts for each individual in the sample. Consequently, one could 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.

Fixed-costs of mobility "matches" transition rates between occupations and between employment and unemployment. Non-monetary benefits "match" the long-run fractions working in each occupations and that cannot be explained by wage differences. The exogenous job breakup rate δ "matches" the observed transition rates from employment to unemployment that is sub-optimal under an unconstrained optimization.

Turning to the parameters describing the approximation functions one can immediately see from equation (4.7) that any intercepts and parameters on variables that are not affected by choice are not identified. It is also clear that parameters on the linear-in-tenure terms cannot be separately identified from non-monetary benefits. However, because of the finite horizon of the decision problem, the value of an additional unit of tenure in occupation j cannot be constant. I set these parameters to zero, allowing me to still interpret the estimates of the B^j-parameters to be "structural".

5. Results

5.1. Parameter Estimates

OLS-regression estimates of the unemployment benefits equation are shown in appendix table 1. Although the model provides a near perfect fit, the parameters are precisely estimated. The evolution of the estimated intercepts over the three time periods correctly reflects the decreasing generosity of the unemployment insurance system in Germany since the mid-80s. In contrast, the results imply that the replacement elasticity has increased since then, possibly absorbing some changes in the system that has benefited individuals with prior strong labor force attachment. On the other hand, the downward adjustment of the replacement rate after one year of unemployment has become larger since the mid-80s.

Appendix tables 2a to 2c show the estimates of the structural parameters in three panels.[29] The first column lists the estimates for the Keane-Wolpin specification with non-monetary benefits and fixed-costs of mobility. The second columns lists the corresponding estimates when two variance components – match heterogeneity and a unit-roots process – are introduced. The bottom of the table shows the outcome of a likelihood-ratio test of the Keane-Wolpin specification against the full model. Evidence strongly favors the latter. They are precisely estimated and highly significant.[30] 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 3.8 percent in the blue-collar occupations, 3.0 percent in the pink-collar occupations, and 6.8 percent in the white-collar occupations. Returns to tenure are very low and generally below 1 percent. Estimates are quite close to those from simple OLS-regressions (not shown in the table) in which the overall

[29] The 110 estimates of the approximation functions are not shown in the table. They are available upon request.

[30] 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 progressively more model features. Most parameters on observables are statistically unchanged across specifications. An exception is 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.

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 blue-collar- or pink-collar occupations but eventually move to white-collar occupations.

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, white-collar occupations experience the strongest trend, and blue-collar occupations experience the weakest trend. The sensitivity to business cycle fluctuations is different across occupations as well. White-collar occupations are the least, and blue-collar 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 statistically and economically significant, and they are quite similar across the two specifications: Employed workers need to pay over 1000 Euro to move into a career they prefer. Unemployed individuals face around 10 percent of the cost of employed workers, possibly reflecting lower opportunity costs of job search, and labor market entrants have negligible mobility costs. The latter result points toward vocational training programs facilitating labor market entry. The corresponding estimates in Keane-Wolpin are significantly higher. For example, they report an estimated cost of almost 4000 USD if a worker is employed. One potential explanation is that they keep different education groups in their sample that are predicted to change into the white-collar occupation when basing decisions on the market wages only. The smaller magnitude of the estimates in comparison to Keane and Wolpin (1997) provides some casual evidence that my estimated cost parameters are "structural" in the sense that the parametric approximation of expected future values absorbs all future expected "exact" values from the Bellman-equations.

It is also important to mention that 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. Transition rates between firms – the dimension of mobility the equilibrium search literature, such as Postel-Vinay and Robin (2002), focuses on – are significantly higher than transition rates between occupations. Consequently my model might miss a large part of earnings fluctuations due to a dimension of mobility that is not explicitly 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 three times as high as the standard deviation of unit roots shocks, and about the same order of magnitude like 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 matches 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.[31]

Estimates of non-monetary benefits are shown at the bottom of appendix table 2b. The estimates need to be interpreted as the non-monetary benefit of a certain alternative relative to the white-collar occupation. Workers are willing to take an annual cut in earnings of 1000 Euro when offered the opportunity to switch from blue- or pink-collar occupations to the white-collar occupation. On the other hand, unemployment provides 100 Euro larger non-monetary benefits per annum than working in the white-collar occupation. Again, these estimates are significantly lower than those reported in Keane and Wolpin (1997), providing further evidence that the Geweke-Keane estimator yields results that are within the ballpark of those found in estimates of full dynamic programming frameworks.

5.2. Model Match

Appendix tables 3 and 4a to 4c together with the appendix figures document the match of the model to the stylized facts listed in section 2. To keep the figures as transparent as possible I only show the match of the full model. The full model matches the wage-experience profiles well, but somewhat over-predicts the growth of variances. The life-cycle profile of the employment share of the white-collar occupation, a stock variable, is matched exceptionally well, especially in the middle of a life-cycle. Other employment shares and the unemployment rates are closely fitted, too. In particular, the model produces the observed re-allocation of labor from the blue- and pink-occupations into the white-occupation. Turning to transition rates reveals a good fit in most dimensions. The only transitions that are not explained very well are those into and out of white-collar occupations and vice versa, which are generally too high later in the life-cycle. Large wage gains to general human capital in white-collar occupations attract individuals with much experience. In the opposite direction, since job breakup rates do not vary across occupations, the model predicts too many transitions between white-collar occupations and unemployment.

Appendix table 3 documents the empirical and predicted relationship between mobility and wage changes. The full model 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 under-predicts the fraction of occupational downgrades that are related to wage increases and over-predicts the fraction of occupational downgrades that are related to wage decreases. An exception is occupational changes between pink-collar and blue-collar occupations, a result explained by the similar wage structures observed in these two occupations. The Keane and Wolpin (1997) framework 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-

[31] 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.

dominantly being driven by iid-transitory shocks to occupational skills which, in the absence of fixed-costs to mobility, generate erratic mobility patterns early in the career.

Appendix table 4a to 4c report 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 Keane-Wolpin model grossly over-predicts earnings mobility, an outcome of transitory shocks representing the only source of residual earnings dynamics.

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 11 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. Consequently, the former group also dominates the latter in terms of wealth, defined as the present value of life-cycle earnings, as depicted in Figure 14. The distributions of the low- and high-earnings types have large overlaps, with low- and high-earners observed in each type-group.

Turning to the analysis of career progression in terms of occupational choice in Figures 12 and 13 reveals that type1-employees are predominantly employed in blue- and pink-collar occupations. With their comparative advantages in these occupations estimated to be very large, they re-allocate toward the white-collar 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 wages than type 1. Although experiencing a stronger re-allocation into the white-collar 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 white-collar occupation which has large returns to experience, and the second group because they quickly re-allocate from the blue- and the pink-collar occupation into the white-collar occupation.

6. Sources of Life-Cycle Career Heterogeneity

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 the present-value of life-cycle earnings, in the following referred to as "wealth". 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. All simulated data replicate the cohort structure of the true data used in the estimation routine.

6.1. Sources of Life-Cycle Income Inequality

I start this section with plotting kernel densities of the present value of life-cycle earnings for the Keane-Wolpin specification and the full model. Results are shown in appendix figure 19, demonstrating one of the key findings of this paper: The Keane- Wolpin model, by abstracting from many potentially important sources of life-cycle earnings dispersion, generates a large role for pre-labor market skill heterogeneity by construction. As clarified by the figure, the kernel density of simulated wealth from the Keane-Wolpin model is multi-modal and is comprised of four very narrow sub-densities – one for each of the four types – that have almost no overlap. As a consequence, while the within-group inequality is small, the across-group inequality is substantial. Once one allows for unit-roots shocks and endogenous match heterogeneity and therefore introduces two sources of within-group inequality, each of the four sub-densities become dispersed enough to make the simulated kernel density approach the empirical density shown in figure 10.

Table 4 reports percentage changes in standard deviations of the wealth distributions when several counterfactual experiments are performed. 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 a random walk and match heterogeneity, 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 full structural 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 heterogeneous 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.[32]

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 73.2 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

[32] It is common, and much faster, to estimate single equation models by GMM.

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 the table. In the Roy-model, 91.5 percent of life-cycle variation is explained by type heterogeneity, a striking result that already has been anticipated by appendix figure 19 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.

Column two shows 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 only 34 percent, 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 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 28 percent.

I also plot the impact of counterfactual experiments on the full distribution of wealth in the full model. Results are provided in figures 15 and 16, with the blue line corresponding to the wealth distribution before conducting the counterfactuals. As anticipated 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 of experience become high enough to prevent individuals to switch.

Figure 16 demonstrates the interesting effect that excluding match heterogeneity exacerbates the role of type heterogeneity. Without match heterogeneity, 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 Keane-Wolpin model. Hence, a substantial part of wealth inequality associated with type-heterogeneity is in fact the outcome of endogenous match-heterogeneity.

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 match-heterogeneity leads to an upward bias 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 bad shocks can pull labor market earnings below unemployment benefits, thus driving a worker into unemployment.

6.2. 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 5 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 increases. 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 anymore insured against by high pre-labor market skills, such as innate abilities.

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 fostering pre-labor market skills are still effective.

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, 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 28 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 comparative advantages or match heterogeneity, earnings mobility increases considerably.

When estimating a proto-typical Roy-model such as considered by Keane and Wolpin (1997) that rules out permanent skill shocks and match heterogeneity – a specification dominating the literature on Discrete Choice Models of post-graduation labor market outcomes – I find the within-type variation to increase to 91.5 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 can significantly influence long-run career outcomes and reduce the incidence of poverty traps. However, the analysis of earnings mobility reveals that pre-labor market skills have a quantitatively important insurance role against particularly bad shocks.

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 explicitly 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 to policies that have never been established before. Third, although I have focused 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, John. M. and David Card (1989): "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*, vol. 57(2), pp. 411-445.

[2] Abowd John, Francis Kramarz and David N. Margolis (1999): "High-Wage Workers and High-Wage Firms." *Econometrica*, 67(2), pp. 251-333.

[3] Adda, Jerome, Christian Dustmann, Costas Meghir and Jean-Marc Robin (2009): "Career Progression and Formal versus On-the-Job Training." Mimeo, University College London.

[4] Adda, Jerome, Monica Costa Dias, Costas Meghir and Barbara Sianesi (2009): "Labor Market Programmes and Labor Market Outcomes." Mimeo, University College London.

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

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

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

[9] Arcidiacono, Peter and Robert Miller (2010): "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," mimeo.

[10] Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay and Jean-Marc Robin (2006): "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation." Mimeo, University of Bristol.

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

[12] Baker, Michael, and Gary 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.

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

[14] Blundell, Richard, Luigi Pistaferri and Ian Preston (2009): "Consumption Inequality and Partial Insurance", *American Economic Review*, 98(5), pp. 1887-1921.

[15] Bontemps, Christian, Jean-Marc Robin and Gerald 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.

[16] Büttner, Thomas, Peter Jacobebbinghaus and Johannes Ludsteck (2008): “Occupational Upgrading and the Business Cycle in West Germany.” Economics Discussion Papers, No 2009-34.

[17] Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006): “Wage Bargaining with On-the-Job Search: Theory and Evidence.” *Econometrica*, vol. 74(2), pp. 323-364.

[18] Cameron, Adrian C. and Pravin K. Trivedi (2005): “Microeconometrics.” Cambridge University Press.

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

[20] Christiano, Lawrence J. and Jonas D. M. Fisher (2000): ”Algorithms for Solving Models with Occasionally Binding Constraints,” *Journal of Economic Dynamics and Control*, vol. 24, pp.1179-1232.

[21] Cunha, Flavio (2007): “A Time to Plant and a Time to Reap.” Mimeo. University of Chicago.

[22] Cunha, Flavio, James J. Heckman (2007): “The Technology of Skill Formation.” *American Economic Review*, vol. 97(2), pp. 31-47.

[23] Cunha, Flavio, James J. Heckman (2007): “The Evolution of Inequality, Heterogeneity and Uncertainty in Labor Earnings in the U.S. economy.” NBERworking papers, w13526.

[24] Cunha, Flavio, James J. Heckman (2008): “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formulation.” *Journal of Human Resources*, vol. 43(4), pp. 738-782.

[25] den Haan, Wouter and Albert Marcet (1990): ”Solving the Stochastic Growth Model by Parameterizing Expectations,” *Journal of Business & Economic Statistics*, Vol. 8, pp. 31-34.

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

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

[28] Duffy, John and Paul D. McNelis (2001): ”Approximating and Simulating the Stochastic Growth Model: Parameterized Expectations, Neural Networks, and the Genetic Algorithm,” *Journal of Economic Dynamic and Control*, Vol. 25, pp.1273-1303.

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

[30] Fitzenberger, Bernd, Aderonke Osikumino and Robert Voelter (2007): "Imputation Rules to improve the EducationVariable in the IAB Employment Subsample", Schmollers Jahrbuch (Journal of the Applied Social Sciences), forthcoming.

[31] Fitzenberger, Bernd and Ralf 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.

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

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

[34] Geweke, John and Michael Keane (2000): "Bayesian Inference for Dynamic Discrete Choice Models without the need for Dynamic Programming," in Roberto Mariano, Til Schuermann and Melvyn J. Weeks (Eds.), Simulation-Based Inference in Econometrics, Cambridge University Press.

[35] Gibbons, Robert and Michael Waldman (1999): "A Theory of Wage and Promotion Dynamics inside Firms." *Quarterly Journal of Economics*, vol. 114(4), pp. 1321-1358.

[36] Gibbons, Robert and Michael Waldman (2006): "Enriching a Theory of Wage and Promotion Dynamics inside Firms." *Journal of Labor Economics*, vol. 24(1), pp. 59-108.

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

[38] Gourieroux, Christian and Alain Monfort (1996): "Simulation-Based Econometric Methods." CORE Lectures, Oxford University Press.

[39] Gourinchas, Pierre-Olivier, and Jonathan A. Parker (2001): "The Empirical Importance of Precautionary Saving", *American Economic Review*, 91(2), pp. 406-412.

[40] Gourinchas, Pierre-Olivier, and Jonathan A. Parker (2002): "Consumption over the Life-Cycle", *Econometrica*, 70(1), pp. 47-89.

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

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

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

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

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

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

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

[48] Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante (2005): "Two Views of Inequality over the Life Cycle", *Journal of the European Economic Association*, 3, pp. 765-775.

[49] Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante (2008): "Insurance and Opportunities: A Welfare Analysis of Labor Market Risk", *Journal of Monetary Economics*, 55(3), pp. 501-525.

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

[51] Heckman, James J., Jora Stixrud and Sergio 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.

[52] Heer, Burkhard and Alfred Maussner (2008): "Dynamic General Equilibrium Modeling - Computational Methods and Applications," 2nd edition, Springer.

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

[54] Hornstein, Andreas, Per Krusell and Gianluca Violante (2006): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." CEPR Discussion Paper 5937.

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

[56] Hotz, V. Joseph, Robert A. Miller, Seth Sanders, and Jeffrey Smith (1994): "A simulation estimator for dynamic models of discrete choice." *The Review of Economic Studies*, 61(2), pp. 265–289.

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

[58] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2006): "Human Capital and Earnings Distribution Dynamics", *Journal of Monetary Economics*, 53(2), pp. 265-290.

[59] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2007): "Sources of Lifetime Inequality", NBER Working Papers 13224.

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

[61] Imai, Susumu, Neelam Jain, and Andrew Ching (2009): "Bayesian estimation of dynamic discrete choice models." *Econometrica*, vol. 77(6), pp. 1865-1899.

[62] Jolivet, Gregory, Fabien Postel-Vinay, and Jean-Marc 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.

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

[64] Judd, Kenneth L. (1998): "Numerical Methods in Economics." The MIT Press, Cambridge, Massachusetts.

[65] Kambourov, Gueorgui and Iurii Manovskii (2009): "Occupational Mobility and Wage Inequality." *Review of Economic Studies*, vol. 76(2), pp. 731-759.

[66] Kaplan, Greg (2007): "Inequality and the Life-Cycle." Mimeo, University of Pennsylvania.

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

[68] Keane, Michael, and Kenneth Wolpin (1997): "The Career Decisions of Young Men", *Journal of Political Economy*, 105(3), pp. 473-522.

[69] Krueger, Dirk, and Fabrizio Perri (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory", *Review of Economic Studies*, 73(1), pp. 163-193.

[70] Krussell, Per and Anthony A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, vol. 106, pp. 867-896.

[71] Lee, Donghoon and Kenneth Wolpin (2006): "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica,* vol. 47, pp. 1-46.

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

[73] Lise, Jeremy. (2007): "On-the-Job Search and Precautionary Savings: Theory and Empirics of Earnings and Wealth Inequality." Mimeo, University College London.

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

[75] Low, Hamish, Costas Meghir and Luigi Pistaferri (forthcoming): "Wage Risk and Employment Risk over the Life Cycle", *American Economic Review*.

[76] MaCurdy, Thomas 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.

[77] Marcet, Albert and David A. Marshall (1992): "Convergence of Approximate Model Solutions to Rational Expectations Equilibria using the Method of Parameterized Expectations," Kellogg Graduate School of Management, Working Paper no. 73.

[78] Marcet, Albert and David A. Marshall (1994): "Solving Non-Linear Rational Expectations Models by Parameterized Expectations: Convergence to Stationary Solutions," mimeo.

[79] Marcet, Albert and Guido Lorenzoni (1999): "The Parameterized Expectations Approach: Some Practical Issues," in Ramon Marimon and Andrew Scott (Eds.), Computational Methods for the Study of Dynamic Economies, Oxford University Press.

[80] Mariano, Roberto, Til Schuermann and Melvyn J. Weeks (2000): "Simulation-Based Inference in Econometrics." Cambridge University Press.

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

[82] Menzio, Guido and Shouyong Shi (2010): "Block Recursive Equilibria for Stochastic Model of Search on the Job." *Journal of Economic Theory*, vol. 145, No. 4, pp. 1453-1494.

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

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

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

[86] Moscarini, Guiseppe and Fabien Postel-Vinay (2009): "Non-stationary Search Equilibrium." Mimeo, Yale University.

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

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

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

[90] Postel-Vinay, Fabien and Jean-Marc Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, vol. 70(6), pp. 2295-2350.

[91] Postel-Vinay, Fabien and Helene Turon (2010): "On-the-job Search, Productivity Shocks, and the Individual Earnings Process." *International Economic Review*, vol. 51, No. 3, pp. 599-629.

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

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

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

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

[96] Storesletten, Kjetil, Chris Telmer, and Amir Yaron (2004a): "Consumption and Risk-Sharing over the Life-Cycle", *Journal of Monetary Economics*, 51, pp. 609-633.

[97] Storesletten, Kjetil, Chris Telmer, and Amir Yaron (2004b): "Cyclical Dynamics in Ideosyncratic Labor Market Risk", *Journal of Political Economy*, 112, pp. 695-717.

[98] Su, Che-Lin and Kenneth L. Judd (2008): "Constrained Optimization Approaches to Estimation of Structural Models," mimeo.

[99] Sullivan, Paul (2010): "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search." *International Economic Review*, 51(1), pp. 298-317.

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

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

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

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

TABLES AND FIGURES

Table 1 – Sample Sizes 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

Table 2 – Mobility and Earnings Changes

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

Table 3 – Earnings 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 – Sources of Earnings Inequality

	Model Specification		Comparison
	Multi-Period Roy Model (Keane&Wolpin)	Full Model With Frictions	Simple Random Walk model
Types	-73.2	-33.8	-61.0
Transitory Shocks	-13.8	-0.3	-0.8
Permanent Shocks	-	-7.1	-12.8
Match Effects	-	-28.1	-
PANEL B: ACROSS TYPE VARIATION IN WEALTH:			
	91.5	41.1	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 – Sources of Earnings Mobility

FULL MODEL WITH FRICTIONS

		Period t+1				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.65	0.22	0.06	0.04	0.03
	Quintile 2	0.24	0.48	0.24	0.05	0.00
Period t	Quintile 3	0.06	0.26	0.43	0.23	0.03
	Quintile 4	0.03	0.04	0.24	0.47	0.21
	Quintile 5	0.03	0.00	0.03	0.21	0.73

NO TYPES

		Period t+1				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.60*	*0.23*	*0.09*	*0.05*	*0.03*
	Quintile 2	*0.25*	*0.41*	*0.25*	*0.08*	*0.01*
Period t	Quintile 3	*0.08*	*0.26*	*0.35*	*0.25*	*0.06*
	Quintile 4	*0.04*	*0.09*	*0.25*	*0.40*	*0.23*
	Quintile 5	*0.03*	*0.01*	*0.06*	*0.23*	*0.67*

NO MATCH HETEROGENEITY

	Quintile 1	*0.58*	*0.26*	*0.09*	*0.04*	*0.04*
	Quintile 2	*0.27*	*0.43*	*0.25*	*0.05*	*0.00*
Period t	Quintile 3	*0.09*	*0.26*	*0.38*	*0.22*	*0.04*
	Quintile 4	*0.03*	*0.05*	*0.23*	*0.43*	*0.25*
	Quintile 5	*0.03*	*0.00*	*0.05*	*0.25*	*0.67*

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

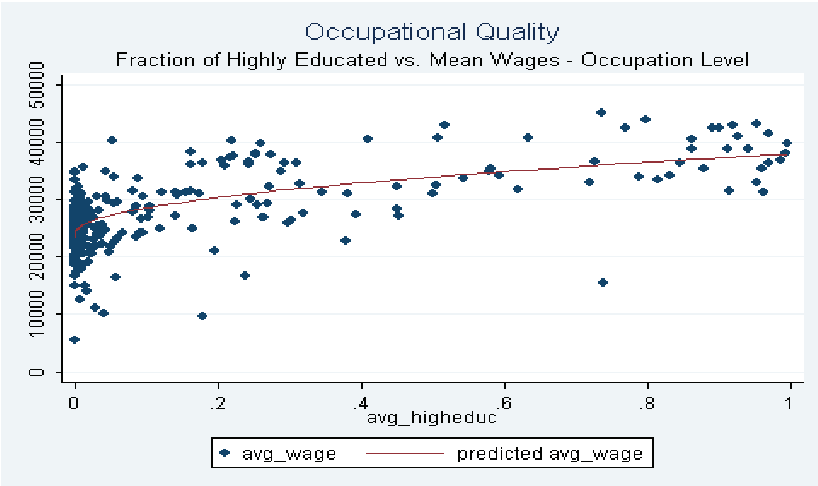


FIGURE 2

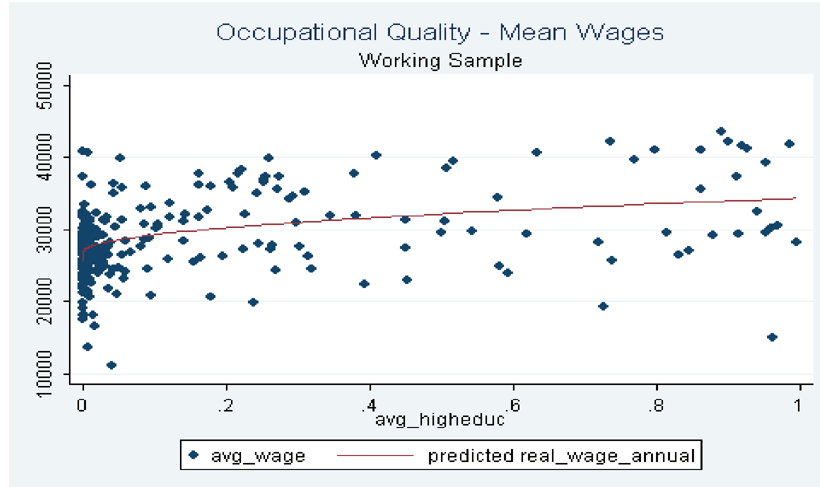


FIGURE 3

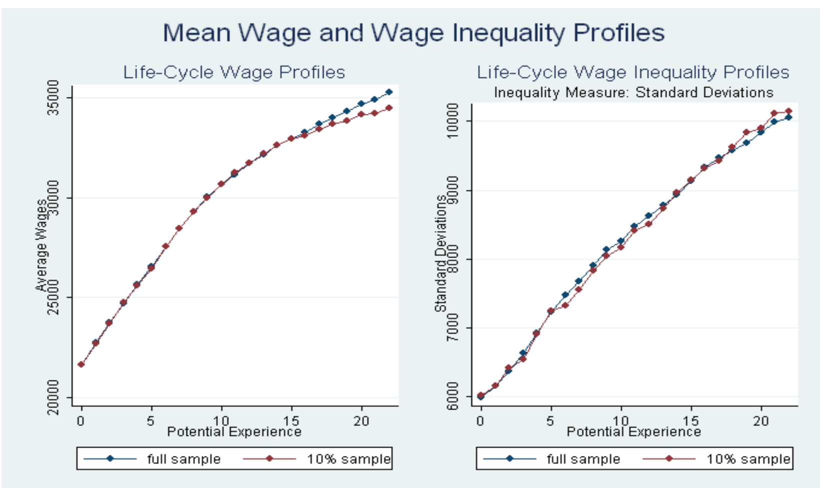


FIGURE 4

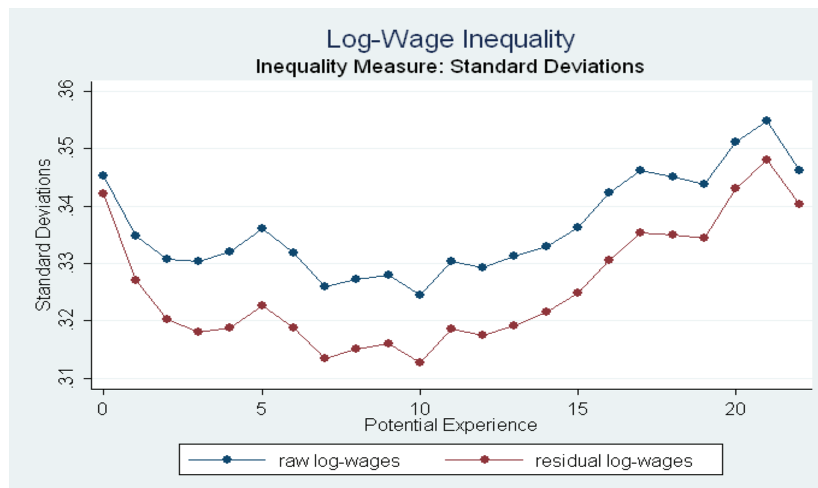


FIGURE 5

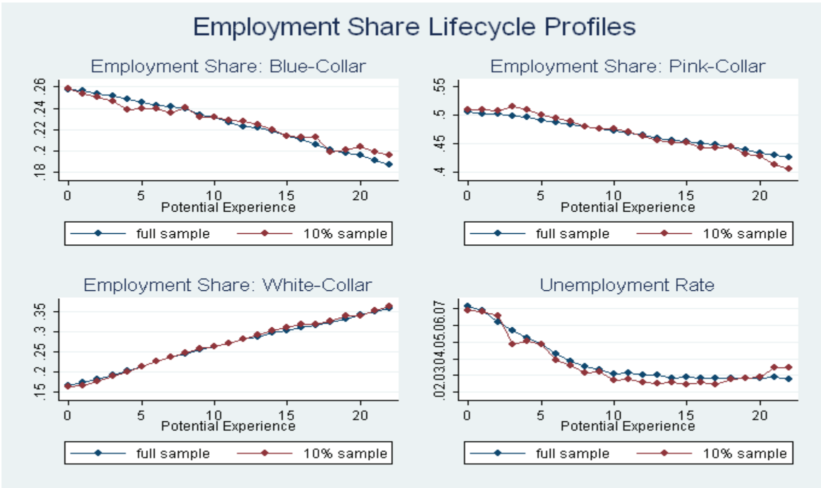


FIGURE 6

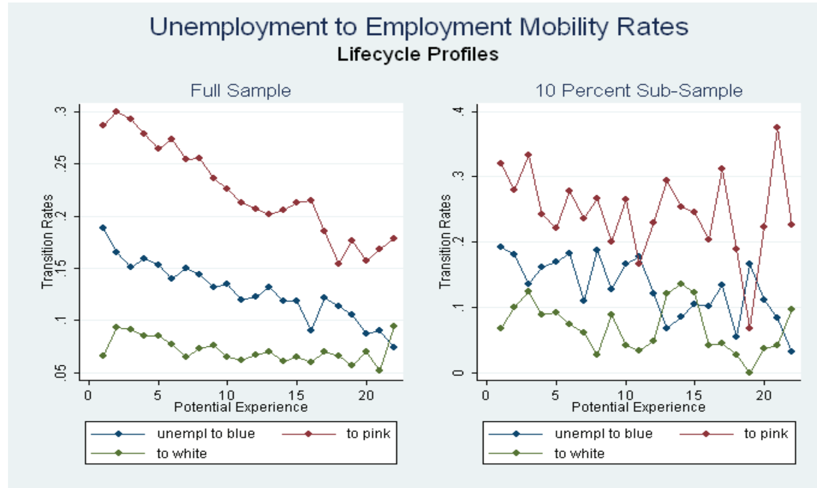


FIGURE 7

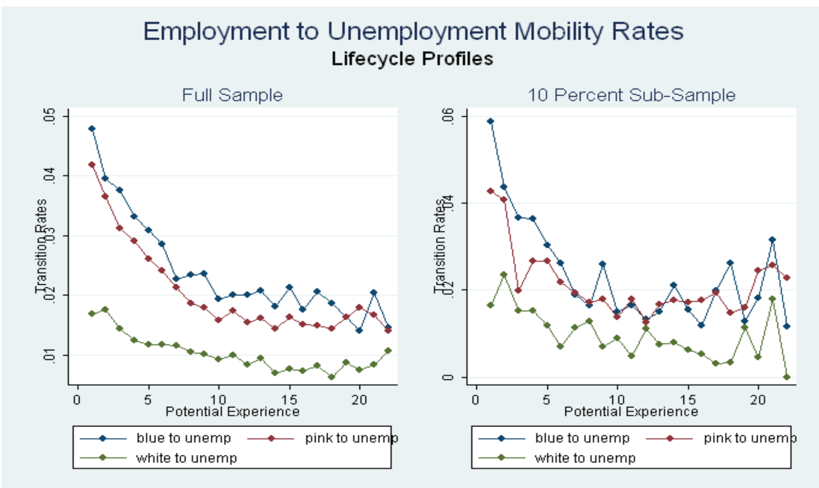


FIGURE 8

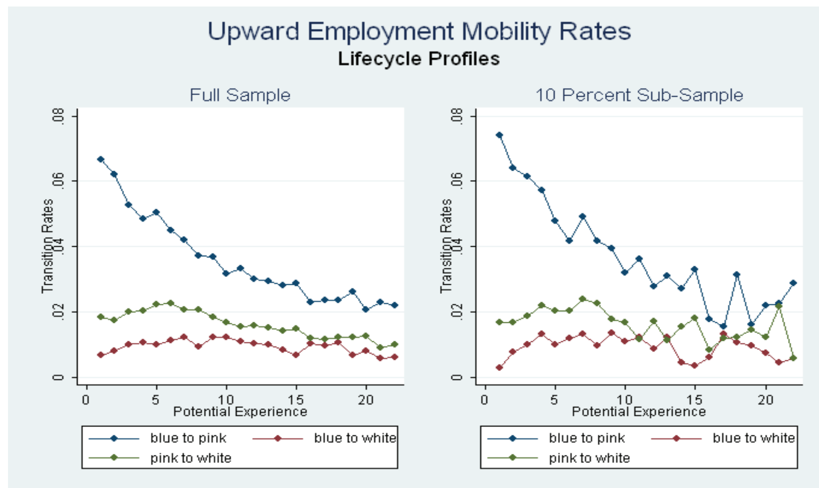


FIGURE 9

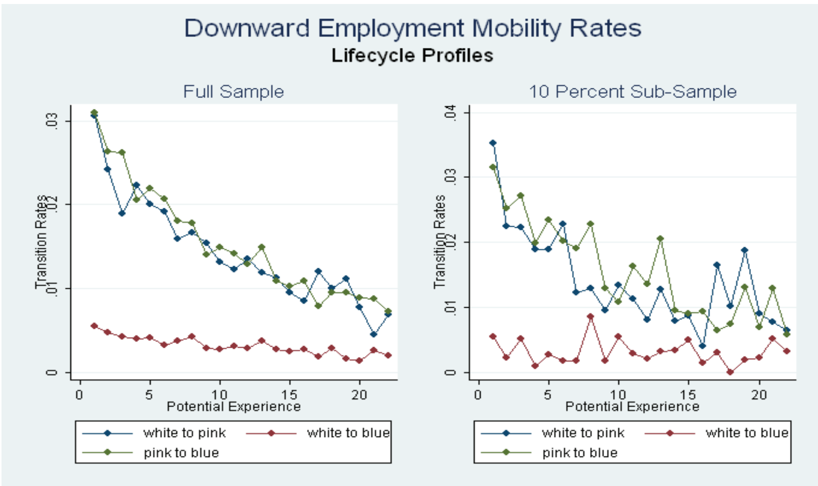


FIGURE 10

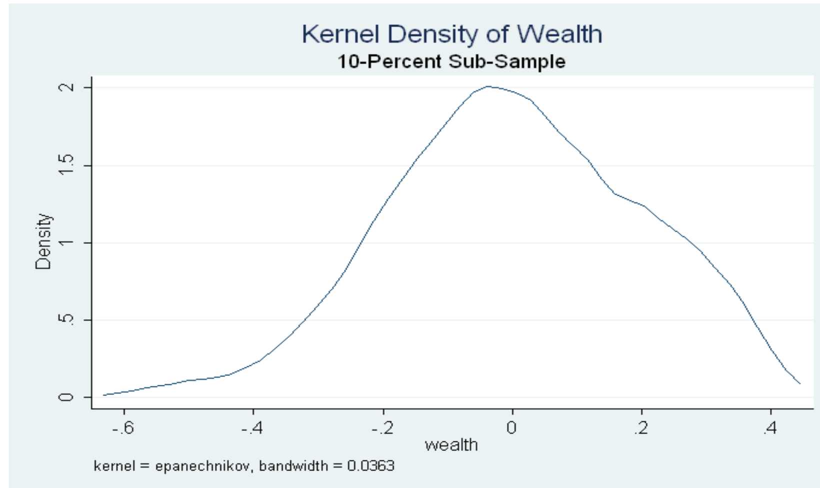


FIGURE 11

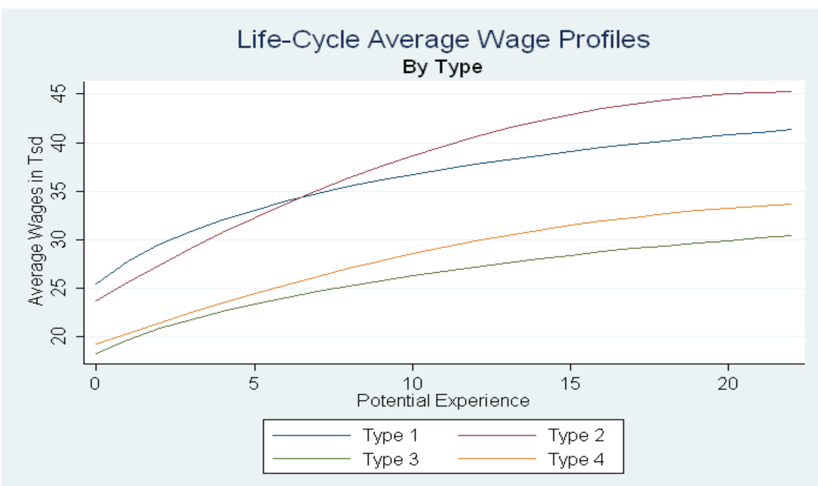


FIGURE 12

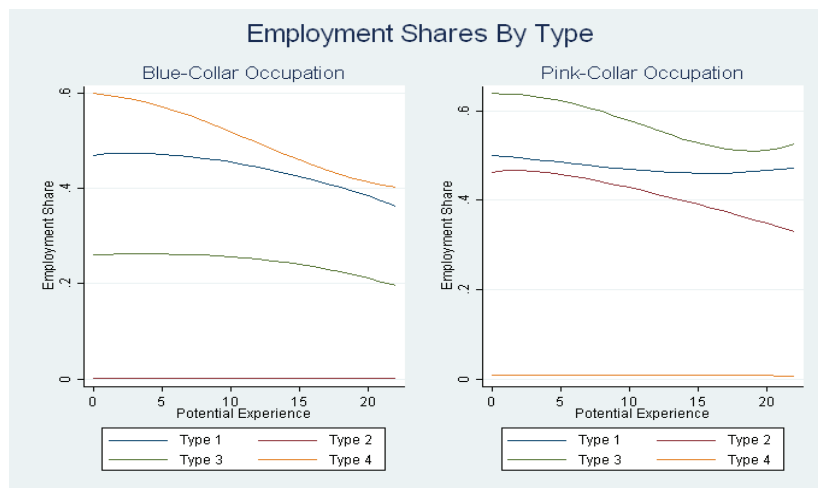


FIGURE 13

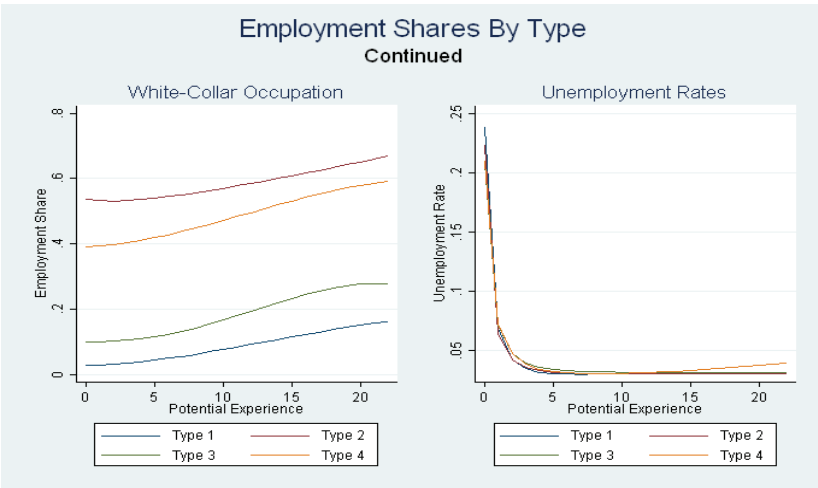


FIGURE 14

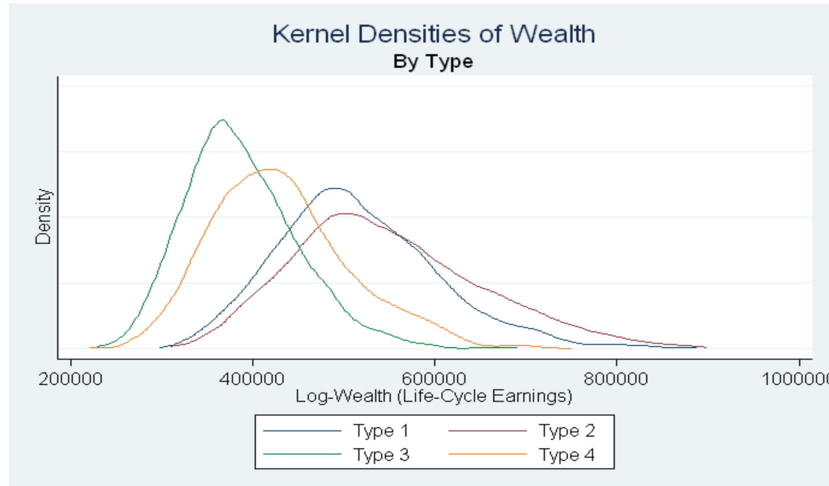


FIGURE 15

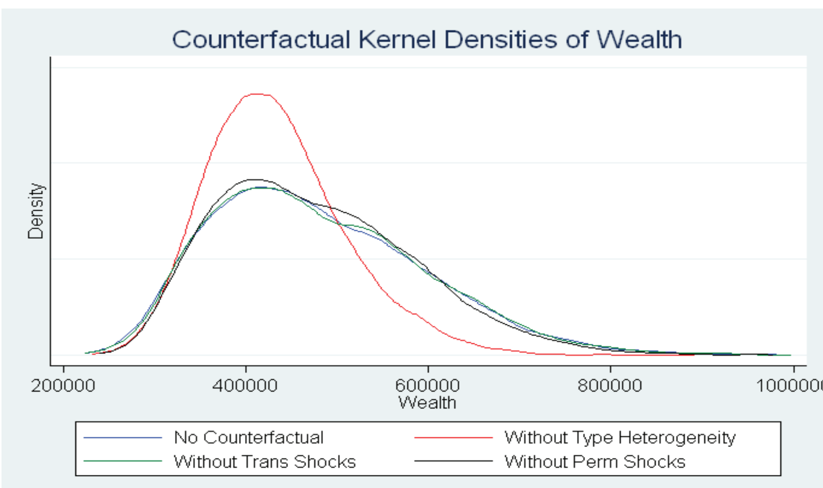
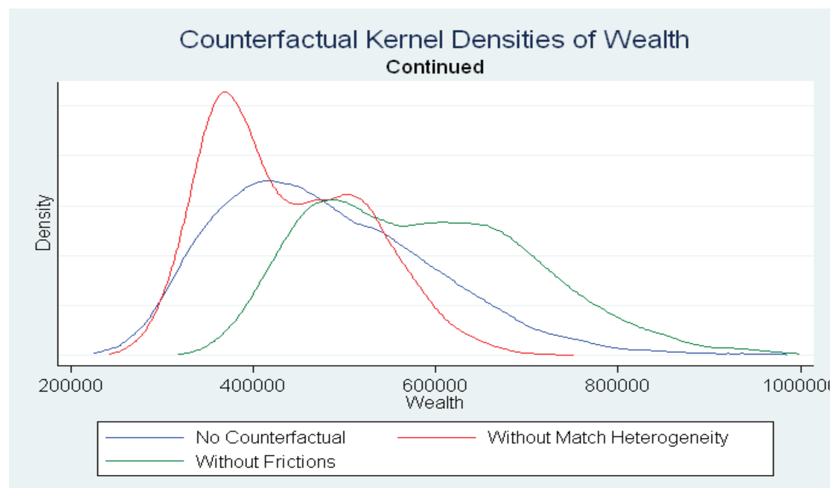


FIGURE 16



APPENDIX TABLES AND FIGURES

Appendix Table 1 – Estimates for Unemployment Benefits Equation

alpha 0 (base amount)	before 1983	9.12 (0.007)	***
	between 1983 and1993	9.02 (0.004)	***
	after 1993	8.91 (0.009)	***
alpha 1 (replacement elasticity AG)	before 1983	0.010 (0.001)	***
	between 1983 and1993	0.019 (0.001)	***
	after 1993	0.032 (0.001)	***
alpha 2 (adjustment of replacement elasticity for AG after 1 year)	before 1983	-0.012 (0.001)	***
	between 1983 and1993	-0.010 (0.001)	***
	after 1993	-0.019 (0.000)	***
	Sample Size (Sample includes only Unemployment Benefits)	36941	
	R-squared	0.9989	

NOTES: This table lists parameter estimates for the model of unemployment benefits. The sample used to estimate these parameters is different from the sample used to estimate the structural Dynamic Discrete Choice Model. It includes all male workers who receive unemployment benefits and are younger than 46 years. *** Significance on 1%-level; ** Significance on 5%-level; * Significance on 10%-level.

Appendix Table 2a – Estimated Structural Parameters

		Keane-Wolpin Model		Full Model	
EQUATION 1 (BLUE-COLLAR OCCUPATION)	Intercept, Type 1	9.102	***	9.471	***
		(0.000)		(0.000)	
	Intercept, Type 2	9.309	***	8.823	***
		(0.000)		(0.000)	
	Intercept, Type 3	9.501	***	9.073	***
		(0.000)		(0.000)	
	Intercept, Type 4	8.603	***	9.244	***
		(0.000)		(0.000)	
	Time Trend	0.084	***	0.208	***
		(0.000)		(0.000)	
	Time Trend^2	0.089	***	0.071	***
		(0.000)		(0.008)	
	Unemployment Rate	-1.680	***	-0.959	***
		(0.000)		(0.052)	
	Experience	-0.049	***	0.038	***
		(0.000)		(0.001)	
	Experience^2	0.003	***	-0.001	***
		(0.000)		(0.000)	
	Tenure	0.149	***	0.004	***
		(0.000)		(0.001)	
	Tenure^2	-0.006	***	0.000	***
		(0.000)		(0.000)	
EQUATION 2 (PINK-COLLAR OCCUPATION)	Intercept, Type 1	8.984	***	9.351	***
		(0.000)		(0.000)	
	Intercept, Type 2	9.201	***	9.321	***
		(0.000)		(0.000)	
	Intercept, Type 3	9.422	***	9.099	***
		(0.000)		(0.000)	
	Intercept, Type 4	8.387	***	8.636	***
		(0.000)		(0.000)	
	Time Trend	-0.084	***	0.273	***
		(0.000)		(0.000)	
	Time Trend^2	0.236	***	0.069	***
		(0.000)		(0.000)	
	Unemployment Rate	-1.547	***	-0.920	***
		(0.000)		(0.031)	
	Experience	-0.033	***	0.030	***
		(0.001)		(0.001)	
	Experience^2	0.002	***	0.000	***
		(0.000)		(0.000)	
	Tenure	0.096	***	0.008	***
		(0.001)		(0.001)	
	Tenure^2	-0.004	***	-0.001	***
		(0.000)		(0.000)	

Appendix Table 2b

EQUATION 3 (WHITE-COLLAR OCCUPATION)	Intercept, Type 1	8.968 (0.000)	***	8.752 (0.020)	***
	Intercept, Type 2	9.525 (0.000)	***	9.089 (0.000)	***
	Intercept, Type 3	9.866 (0.000)	***	8.567 (0.018)	***
	Intercept, Type 4	6.586 (0.000)	***	8.805 (0.018)	***
	Time Trend	0.098 (0.000)	***	0.380 (0.000)	***
	Time Trend^2	-0.003 (0.000)	***	0.073 (0.000)	***
	Unemployment Rate	-1.180 (0.000)	***	-0.677 (0.048)	***
	Experience	0.012 (0.000)	***	0.068 (0.001)	***
	Experience^2	0.003 (0.000)	***	-0.002 (0.000)	***
	Tenure	0.099 (0.000)	***	0.002 (0.000)	***
	Tenure^2	-0.006 (0.000)	***	0.000 (0.000)	***
FIXED-COSTS OF MOBILITY	Labor Market Entrants	35.7 (0.732)	***	38.1 (0.818)	***
	Unemployed	121.2 (15.282)	***	127.4 (18.082)	***
	Employed	1061.8 (18.532)	***	1061.2 (19.576)	***
NON-MONETARY BENEFITS RELATIVE TO WHITE-COLLAR OCCUPATION	Labor Market Entrants	100.5 (1.232)	***	113.7 (1.091)	***
	Unemployed	-1086.7 (63.520)	***	-1086.8 (80.550)	***
	Employed	-1030.7 (19.134)	***	-1036.4 (25.070)	***

Appendix Table 2c

VARIANCE COMPONENTS (IN STANDARD DEVIATIONS)	Trans.Shock, Blue Collar	0.296 (0.000)	***	0.112 (0.001)	***
	Trans. Shock, Pink Collar	0.236 (0.000)	***	0.103 (0.000)	***
	Trans Shock, White Collar	0.205 (0.000)	***	0.105 (0.000)	***
	Perm Shock, General Skills	-		0.034 (0.000)	***
	Match Effects	-		0.101 (0.000)	***
TYPES AND TYPE PROPORTIONS	Fraction of Type 1	0.138	***	0.241	***
	Fraction of Type 2	0.392	***	0.247	***
	Fraction of Type 3	0.397	***	0.267	***
	Fraction of Type 4	0.072	***	0.245	***
	NO OF OBSERVATIONS		85,319		
	LOG-LIKELIHOOD TEST AGAINST KEANE-WOLPIN (P-VALUE)		0.000		

NOTES: This table lists parameter estimates from the Keane-Wolpin model with non-monetary benefits and fixed costs of mobility. *** Significance on 1%-level; ** Significance on 5%-level; * Significance on 10%-level.

Appendix Table 3 – Model Match to Relationship between Mobility and Earnings Changes

		Wage Increase			Wage Decrease		
		Data	Keane-Wolpin	Full Model	Data	Keane-Wolpin	Full Model
Upward Mobility	Blue to Pink	0.47	0.46	0.62	0.35	0.47	0.34
	Blue to White	0.63	0.37	0.63	0.22	0.56	0.28
	Pink to White	0.61	0.42	0.65	0.24	0.54	0.28
Downward Mobility	White to Pink	0.49	0.59	0.35	0.32	0.35	0.60
	White to Blue	0.58	0.63	0.41	0.26	0.33	0.52
	Pink to Blue	0.49	0.51	0.49	0.35	0.43	0.41

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 "blue-collar", "pink-collar" and "white-collar" is described in the text. A discrete wage change is defined as a wage change that is larger than 2 percent of the average wage in the sample.

Appendix Table 4a – Model Match to Earnings Mobility

		DATA					ROY-MODEL (KEANE & WOLPIN)				
		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	0.80	0.14	0.03	0.02	0.01	*0.59*	*0.21*	*0.12*	*0.05*	*0.02*
	Quintile 2	0.15	0.65	0.16	0.03	0.01	*0.21*	*0.31*	*0.24*	*0.15*	*0.08*
Period t	Quintile 3	0.03	0.17	0.61	0.16	0.02	*0.12*	*0.24*	*0.25*	*0.22*	*0.18*
	Quintile 4	0.02	0.02	0.18	0.64	0.14	*0.05*	*0.16*	*0.22*	*0.28*	*0.30*
	Quintile 5	0.01	0.01	0.02	0.15	0.82	*0.02*	*0.09*	*0.18*	*0.29*	*0.42*

		FULL MODEL				
	Quintile 1	0.65	0.22	0.06	0.04	0.03
	Quintile 2	0.24	0.48	0.24	0.05	0.00
Period t	Quintile 3	0.06	0.26	0.43	0.23	0.03
	Quintile 4	0.03	0.04	0.24	0.47	0.21
	Quintile 5	0.03	0.00	0.03	0.21	0.73

Appendix Table 4b

		DATA					ROY-MODEL (KEANE & WOLPIN)				
		Period t+5					Period t+5				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.62*	*0.22*	*0.09*	*0.04*	*0.03*	*0.57*	*0.22*	*0.12*	*0.06*	*0.02*
	Quintile 2	*0.25*	*0.41*	*0.22*	*0.10*	*0.03*	*0.22*	*0.24*	*0.26*	*0.17*	*0.10*
Period t	Quintile 3	*0.09*	*0.25*	*0.37*	*0.22*	*0.08*	*0.11*	*0.29*	*0.23*	*0.20*	*0.18*
	Quintile 4	*0.05*	*0.09*	*0.25*	*0.41*	*0.21*	*0.05*	*0.21*	*0.20*	*0.25*	*0.28*
	Quintile 5	*0.03*	*0.03*	*0.07*	*0.23*	*0.65*	*0.03*	*0.09*	*0.18*	*0.30*	*0.40*

		FULL MODEL				
	Quintile 1	0.53	0.25	0.11	0.06	0.05
	Quintile 2	0.26	0.38	0.24	0.09	0.02
Period t	Quintile 3	0.10	0.25	0.34	0.24	0.08
	Quintile 4	0.05	0.09	0.24	0.38	0.24
	Quintile 5	0.04	0.02	0.07	0.23	0.63

Appendix Table 4c

DATA

		Period t+10				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.49*	*0.25*	*0.13*	*0.07*	*0.05*
	Quintile 2	*0.26*	*0.31*	*0.23*	*0.14*	*0.06*
Period t	Quintile 3	*0.13*	*0.24*	*0.28*	*0.23*	*0.12*
	Quintile 4	*0.07*	*0.13*	*0.24*	*0.31*	*0.25*
	Quintile 5	*0.04*	*0.05*	*0.11*	*0.25*	*0.55*

ROY-MODEL (KEANE & WOLPIN)

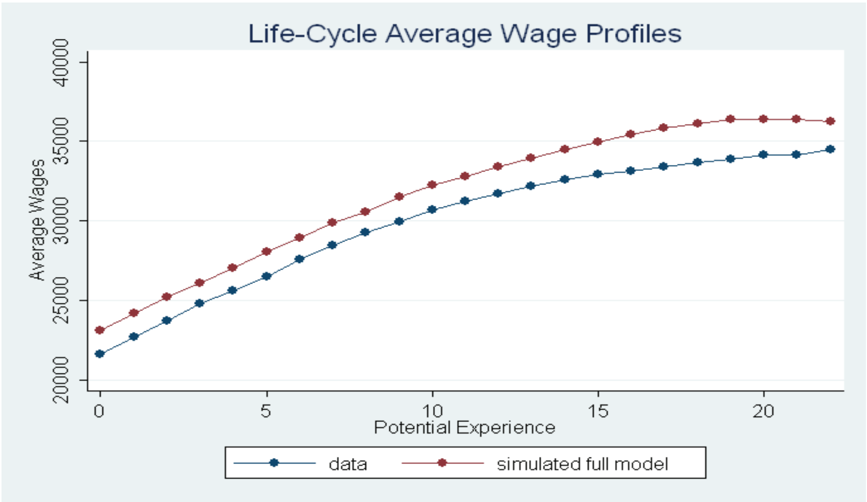
		Period t+10				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	*0.55*	*0.22*	*0.14*	*0.07*	*0.03*
	Quintile 2	*0.19*	*0.29*	*0.24*	*0.17*	*0.11*
Period t	Quintile 3	*0.11*	*0.22*	*0.23*	*0.23*	*0.22*
	Quintile 4	*0.06*	*0.15*	*0.21*	*0.27*	*0.31*
	Quintile 5	*0.03*	*0.10*	*0.19*	*0.30*	*0.38*

FULL MODEL

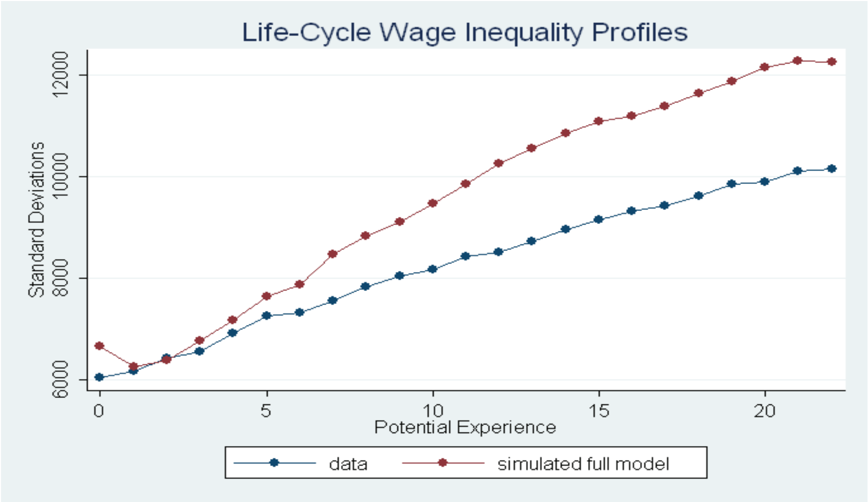
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Quintile 1	0.44	0.26	0.15	0.09	0.06
	Quintile 2	0.26	0.32	0.24	0.13	0.05
Period t	Quintile 3	0.13	0.23	0.27	0.23	0.13
	Quintile 4	0.07	0.12	0.23	0.32	0.26
	Quintile 5	0.05	0.05	0.12	0.25	0.54

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.

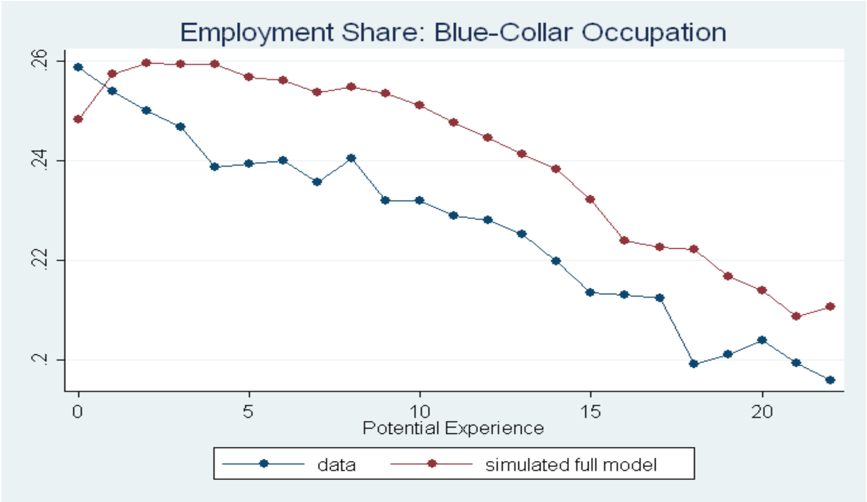
APPENDIX FIGURE 1



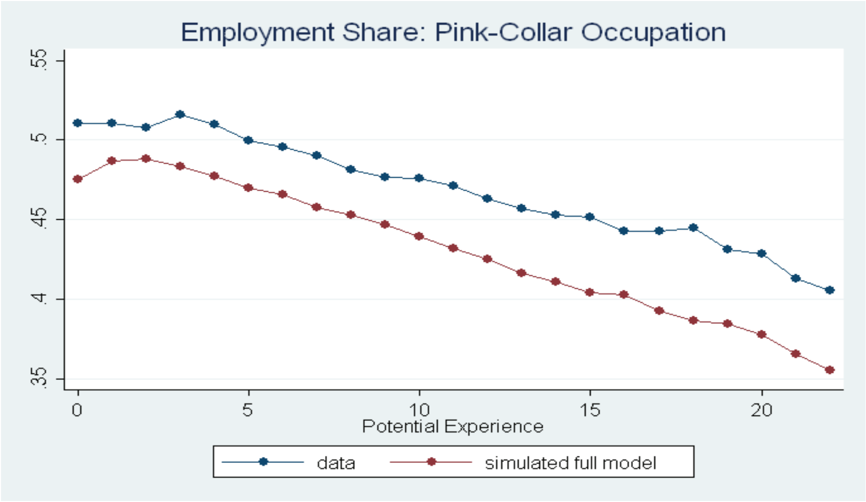
APPENDIX FIGURE 2



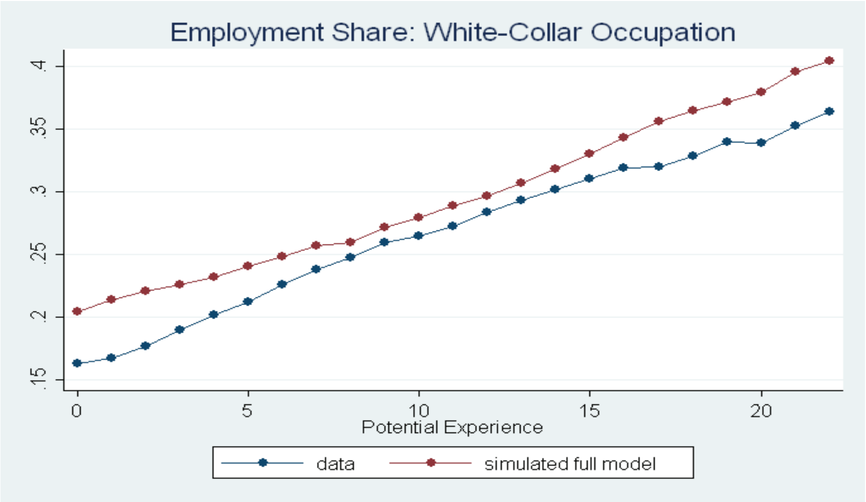
APPENDIX FIGURE 3



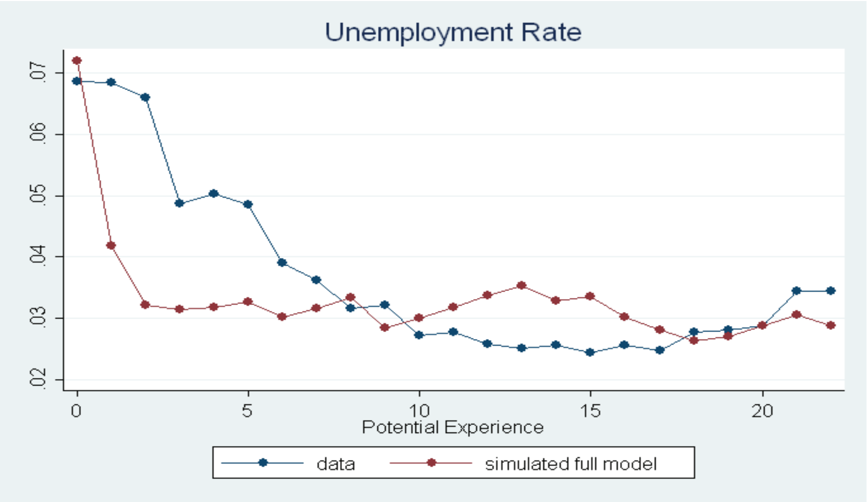
APPENDIX FIGURE 4



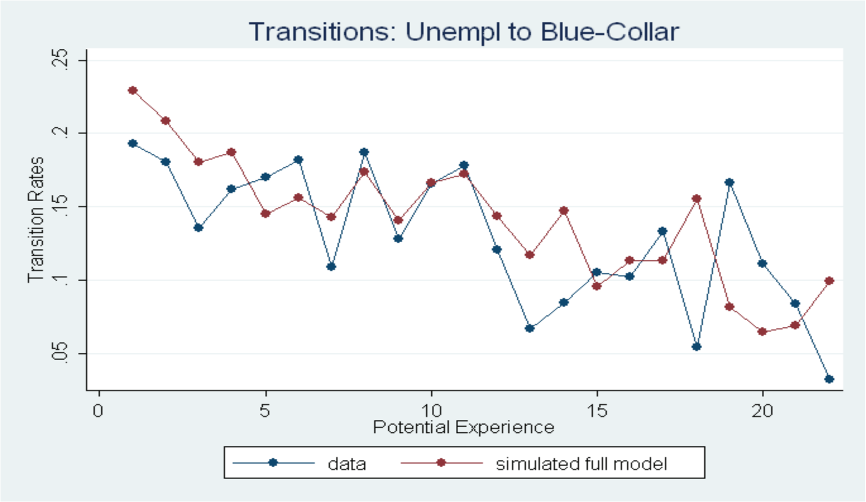
APPENDIX FIGURE 5



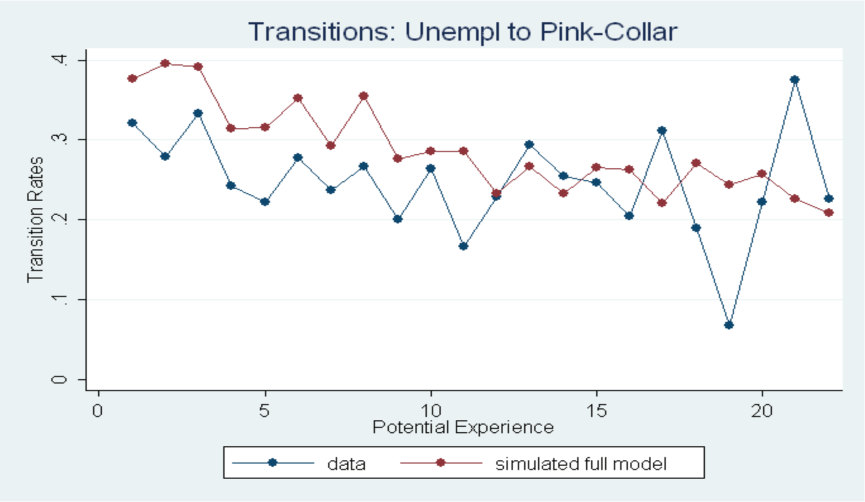
APPENDIX FIGURE 6



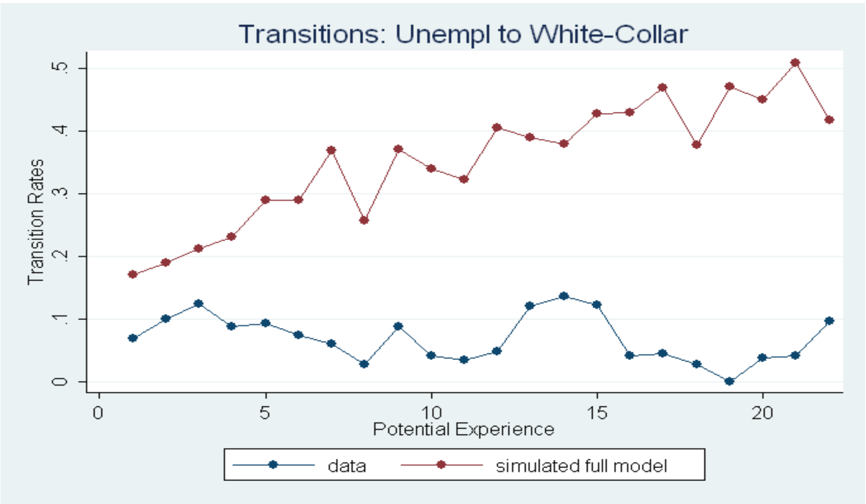
APPENDIX FIGURE 7



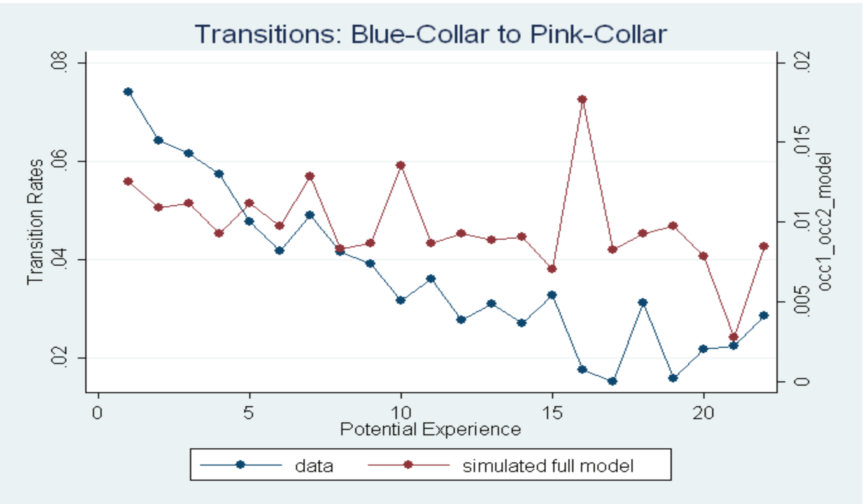
APPENDIX FIGURE 8



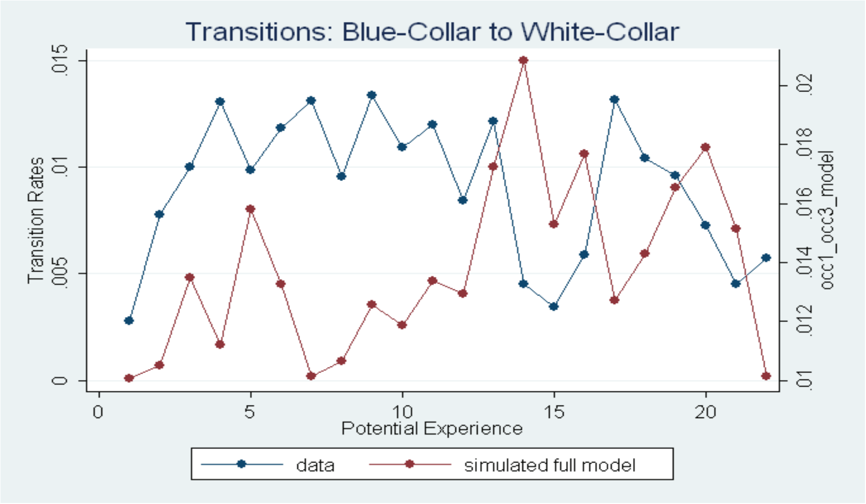
APPENDIX FIGURE 9



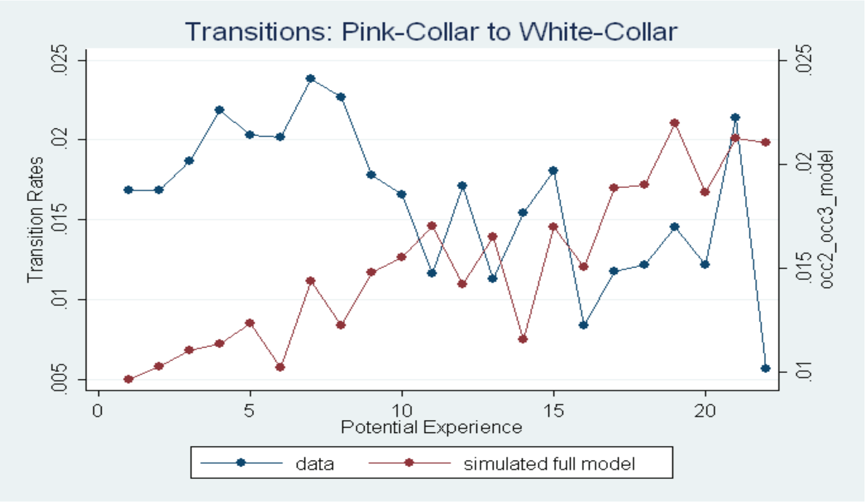
APPENDIX FIGURE 10



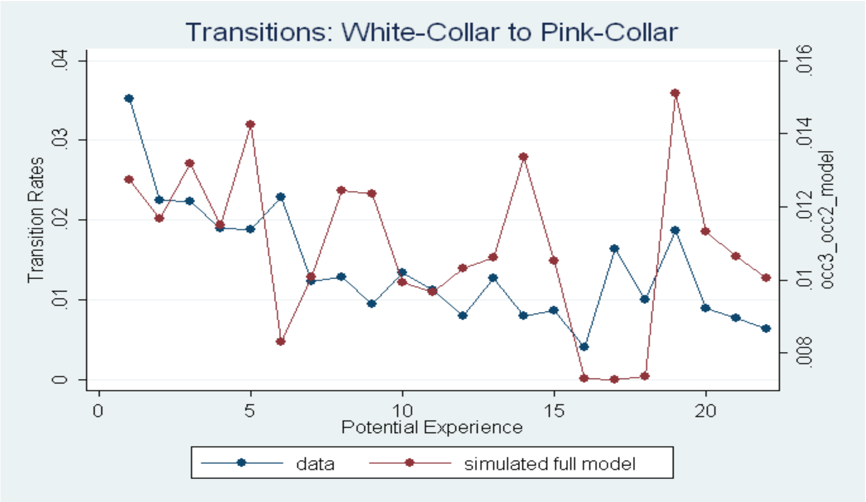
APPENDIX FIGURE 11



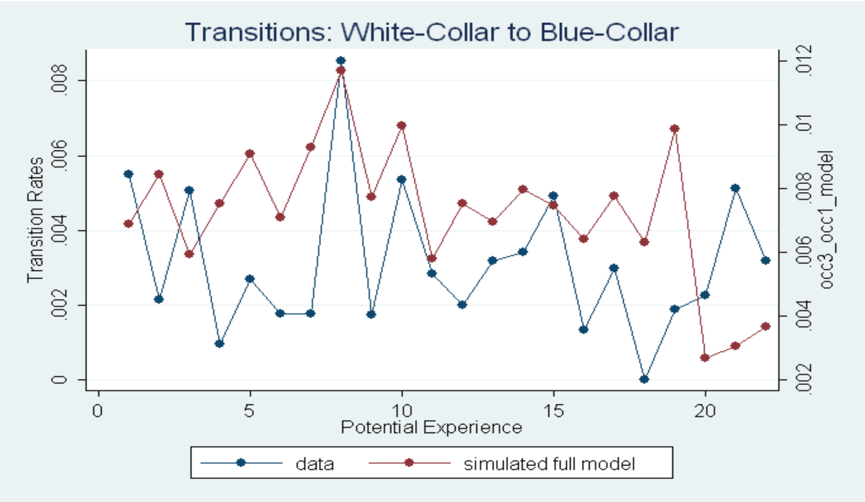
APPENDIX FIGURE 12



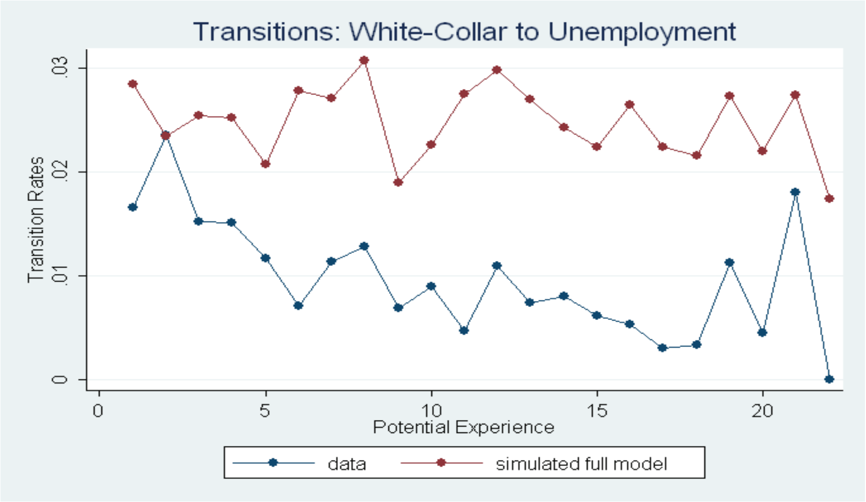
APPENDIX FIGURE 13



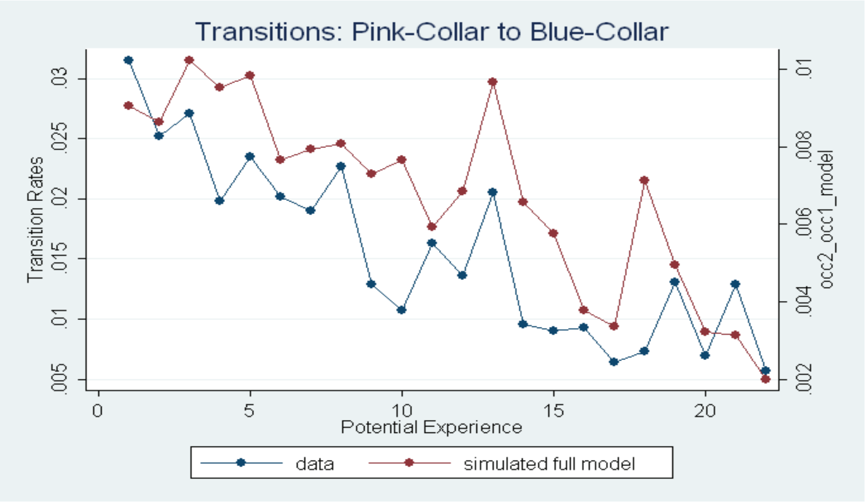
APPENDIX FIGURE 14



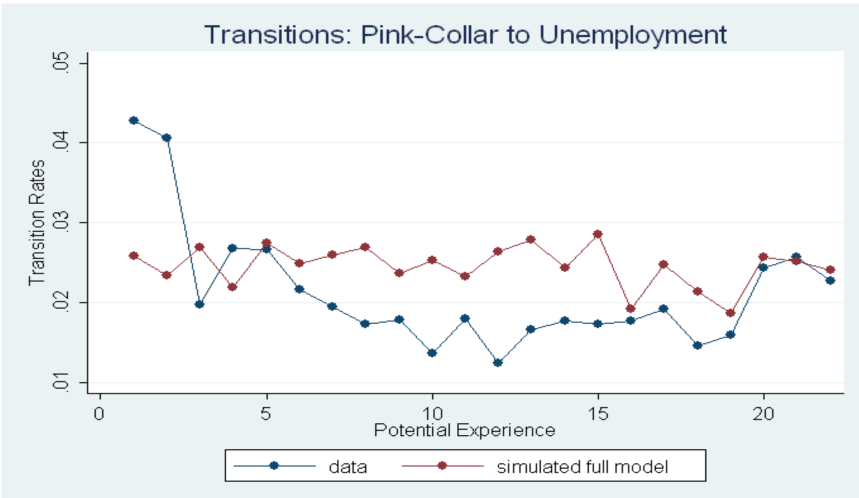
APPENDIX FIGURE 15



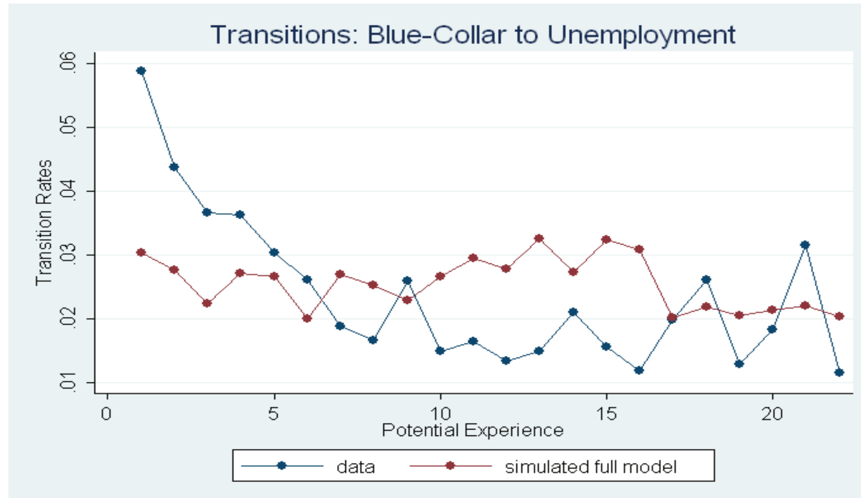
APPENDIX FIGURE 16



APPENDIX FIGURE 17



APPENDIX FIGURE 18



APPENDIX FIGURE 19