

Learning and Life Cycle Patterns of Occupational Transitions*

Aspen Gorry

Nicholas Trachter

UC Santa Cruz

EIEF

aspen@ucsc.edu

nicholas.trachter@eief.it

October 31, 2011

Abstract

This paper develops a simple life cycle model of occupational choices based on workers learning about their type and sorting themselves to the best match. Documenting life cycle patterns of occupational choices using data from the NLSY79 provides evidence for key predictions from the model. First, initial characteristics are predictive of future patterns of occupational switching. Also, the average time to the first occupational switch is longer than the time to the second switch. Finally, the data reveal a high number of occupational switches as almost 40% of high school graduates switch between white and blue collar occupations more than once between the ages of 19 and 28.

Keywords: Learning, Occupational Mobility, Labor Markets, Life Cycle.

JEL Codes: E24, J24, J31, J62.

*In progress and preliminary, comments welcome. We greatly appreciate research assistance from Timothy Elser and Matthew Pecenco. All mistakes are our own.

1 Introduction

Worker turnover is a primary feature in understanding individual labor market dynamics (see [Topel and Ward \(1992\)](#)). Occupational choices are some of the most important decisions facing young workers. Recent research shows that much of the human capital gained through experience occurs at the level of worker’s occupation or industry.¹ Moreover, occupational mobility can account for a large amount of observed wage inequality (see [Kambourov and Manovskii \(2009a\)](#)).

Given that so much human capital accumulates at the occupational level it is striking that occupational mobility has been increasing in the United States between 1968 and 1997 ([Kambourov and Manovskii \(2008\)](#)). The current literature has documented patterns of occupational mobility in the cross section of workers. This paper starts by documenting life cycle patterns of occupational mobility from the NLSY79. By looking at high school educated worker’s occupational choices between the ages of 19 and 28 more detailed facts about individual level occupational patterns emerge beyond the well known fact that occupational switches decline with age. The first finding is that initial characteristics are predictive of both the initial occupational choice and future decisions to switch.

Next, we find that there a large group of the population that switches occupations frequently during their first 10 years. While Evidence on quantity and timing of switches. 42% of people do not switch between blue and white collar during these 10 years while another 19% switch exactly once. However, almost 40% of the sample switches between blue and white collar more than one time during their first 10 years. This result is similar to that found in [James \(2011\)](#). Using NLSY97 data he finds that a large portion of workers switch back to their previous occupation.

Finally, we examine the timing work worker’s occupational switches and find that for

¹Papers by [Neal \(1995\)](#) and [Parent \(2000\)](#) argue that human capital is largely based at the industry level rather than being firm specific while more recent research by [Kambourov and Manovskii \(2009b\)](#) argue that human capital is based at the occupational level. [Poletaev and Robinson \(2008\)](#) and [Gathmann and Schönberg \(2010\)](#) have shown skills to be task specific. Even under this view skills show strong correlation within occupational categories.

workers who switch more than once the average time to their first switch is on average longer than the average time to their second switch. This result is striking in contrast to the cross sectional result that workers average rate of switching declines with age.

The findings about initial information being important and the time to the worker's second switch being on average shorter than the first is consistent with models where workers learn about their abilities such as Papageorgiou (2010). We construct life cycle version of Johnson's (1978) model where individuals choose their initial occupation based on their prior beliefs about their ability type and learn about their ability by observing their production on their current job. The workers initial belief about their type will cause them to make an initial occupational choice. After observing output each period they update their beliefs and can choose to switch occupations. Despite one occupation being riskier than the other, workers may still choose the occupation where learning is slower as they may be avoiding wage risk across sectors or gain a wage premium for the lower learning sector based on their beliefs.

Using the predicted probability of workers initially choosing to enter the white collar sector based on observable characteristics from the NLSY79, we are able to generate predictions about future occupational transitions. First, we are able to confirm the prediction from the model that for two workers in the same sectors with different beliefs, the worker with the belief closest to the other sector will on average switch sooner than the worker with the belief that is further away. Second, we find that conditional on switching once the time to a second switch is less than the initial time. This provides further evidence for the mechanism in the paper as workers who switch are likely to be closer to the barrier and hence more likely to switch.

Finally, we evaluate the model's ability to generate the observed quantity of switching. The baseline model has difficulty in generating the number of workers that switch occupations multiple times as found in the data. By extending the model to include individual level wage shocks across occupations, the model is able to generate much more switching.

The model in this paper is most closely related to Papageorgiou (2010), Johnson (1978),

and Miller (1984). We build a learning model into an explicit life cycle framework. Early studies such as Johnson (1978) and Miller (1984) emphasize uncertainty and imperfect information to explain worker’s occupational sequencing decisions. The implication of these models is that young workers should initially choose riskier professions for given expected returns². In models of learning, the higher uncertainty in some occupations allows workers to learn about where they are more productive more rapidly. Jovanovic and Nyarko (1997) criticize this method of learning arguing that in observed occupational sequences people learn by completing simple tasks first. Learning should create job ladders instead of transitions from more to less risky tasks.

Other related papers on learning and occupational choices include McCall (1990) who demonstrates the presence of occupation specific learning by showing that job switches within an occupation lead to longer job durations than switches across occupations, Neal (1999) who constructs a model of occupation and job specific matches, and Antonovics and Golan (2010) who extend the standard learning literature by allowing workers to choose the rate at which they learn.

Observed occupational transitions are actually much more complex than described by either of these types of models. While simple learning models are useful for understanding occupational choices in isolation, these frameworks relate to the literature that models individual career patterns using a discrete choice framework following Keane and Wolpin (1997). Recent papers in this literature have extended Keane and Wolpin (1997) by including richer features such as occupational and match heterogeneity and shocks to skills in Hoffmann (2010) and inter-firm mobility and firm specific human capital in Sullivan (2010).³ Worker’s choose a variety of initial occupations, switch occupations frequently, and move in both directions from riskier to less risky occupations. To rationalize these observations, we extend

²This result was first demonstrated by Johnson (1978) who shows that risk neutral agents will choose occupations with higher earnings variance in the first period in life. Miller (1984) generalizes the model to allow for different occupation types and provides empirical support for job switching toward certain types of risk.

³Recent papers on the dynamics of occupational mobility include Keane and Wolpin (1997) and Groes et al. (2010).

previous learning models of occupational choice by including risk averse workers and show that risk aversion can account for the observed patterns of occupational choices in NLSY79 data.

Section 2 documents life cycle patterns of occupational data using data from NLSY79. Section 3 presents the simple learning model. Section 4 describes the parameterization of the model and Section 5 presents the results. Section 6 concludes.

2 Life Cycle Occupational Switches: Evidence

This paper uses data from the National Longitudinal Survey of Youth 1979 (NLSY79) to document the life cycle patterns of occupational choices. The NLSY79 is a nationally representative longitudinal survey conducted by the Bureau of Labor Statistics that samples 12,686 individuals who were between the ages of 14 and 22 years old when first surveyed in 1979. The individuals continued to be surveyed every year until 1994 when the survey switched to every two years. NLSY79 provides a rich set of panel data for tracking worker's career outcomes.

The primary analysis in this paper will focus on the annually reported CPS job for high school graduates from the age of 19-28. The data are restricted to before 1994 so that annual observations of CPS job and wage data are available. Restricting the sample to people who were at most 19 years old when first interviewed in 1979 and have a high school degree as the highest degree ever received during the entire sample leaves 4,180 individuals in the sample. These restrictions ensure a group who is uniformly in the labor force between the ages of 19 and 28.

To analyze individual occupational choices the annual CPS job for each worker is used to construct the worker's choice of occupation⁴. By focusing on broad occupation categories the analysis avoids problems in the NLSY79 with coding errors of occupations at finer levels

⁴We follow [Keane and Wolpin \(1997\)](#) in defining blue and white collar occupations using one digit codes. Blue collar occupations include craftsmen, foremen and kindred, farm laborers and foremen, and service workers. White collar occupations are professional, technical, and kindred, managers, officials, and proprietors, sales workers, farmers and farm managers, and clerical and kindred.

of detail that are discussed in [Keane and Wolpin \(1997\)](#). Using the CPS occupation for each year, the individual’s occupational transitions can be tracked at a yearly frequency⁵.

To understand workers initial job choices, [Table 1](#) presents summary statistics for the variables used in the analysis by the workers initial occupational choice along with the total number of observations for each variable. Standard deviations are reported for the non-categorical variables. The variables for mother’s education and father’s education record the number of years of education earned by the mother and father respectively. The variables mother’s and father’s main occupation are classified into blue and white collar (taking the value of 1 if white collar). There are also categorical variables for sex (male), geography (urban), race (white and black with a base case of hispanic) and family poverty (takes a value of 1 if the family is ever marked as being in poverty). Class percentile gives the respondents percentile rank in the last year they attended school that is available in the 1981 survey. Finally, using the Armed Forces Qualification Test (AFQT) subject scores, the respondent’s percentile in each subsection is constructed. Given the restriction to high school graduates, about 35% of the sample’s first occupation is white collar. Individuals that choose an initial occupation in the white collar sector tend to have parents with more education and who work in white collar occupations, be female, white, live in a city, not have been in poverty, and have a higher class rank and AFQT score. It should be noted that the number of observations drop substantially for the parental occupation variables and class percentile. To deal with these issues the parental occupation variables are dropped in some specifications and the missing values of class percentile are imputed using the other variables in the table.

⁵This method ignores the possibility of multiple switches within the year. This is unlikely to be a major issue as broad occupational categories are used. Moreover, recoding occupational switch variables to account for multiple switches within a year does not change any of the results. This second measure may overstate actual switches as it also includes secondary jobs.

2.1 Initial Occupational Choice

The first stage of the analysis evaluates how useful initial characteristics (known at the time of first occupational choice) are in predicting individual’s occupational decisions. **Table 2** reports the results from a probit regression of the initial observable characteristics on the worker initially choosing the white collar occupation. In the first specification the variables for mother’s and father’s main occupation are dropped. The second specification includes these additional variables. The values in the table are reported as marginal effects. In the first specification, an additional year of Father’s schooling generates a 1.4% increase in the probability of choosing white collar. Being male and white reduce the probability of choosing white collar by 29% and 8% respectively. Living in a city increases the probability of choosing white collar by 11% while having experienced poverty decreases the probability by 3.9%. Finally, a 10% increase in class and AFQT test percentile increase the probability of choosing a white collar initial occupation by 3.6% and 1.4% respectively. Mother’s education is not significant. In the second specification, all variables have the same sign and all remain significant except for the AFQT percentile. Additionally, if the respondent’s mother works in a white collar occupation they have a 6.8% higher chance of working in white collar. Father’s main occupation is not significant.

While the results of the probit on initial choice are of interest, a more interesting question is if the predicted initial choices of workers have predictive power about workers future occupational choices. **Table 3** reports the results from a second probit regress of whether individual’s ever switch occupations based on the fitted probability of their choosing a white collar occupation initially based on the first probit regression. The first two columns show results from the first specification where the fitted probabilities are constructed from the probit where parental occupation variables have been dropped. The results are broken down by the individual’s initial occupational choice. The table shows that workers with a 10% higher imputed probability of choosing a white collar initial occupation are 6.8% less likely to ever switch occupations if they start in white collar. Similarly, the second column shows that for workers who start in blue collar are 4.5% more likely to switch occupations if their

fitted probability of choosing white collar is 10% higher. The final two columns replicate the results for the specification including parental occupational variables. The magnitudes and significance of the findings are similar.

Finally, [Table 4](#) examines if the fitted probability of initially choosing white collar is informative about the timing of switches for those workers that switch occupations. The table shows the results of regressions of the timing of the first occupational switch on the fitted probability of choosing white collar. Separate regressions are run by observed initial occupational choice. In the first specification the probabilities are constructed from the probit without parental occupation variables. The regressions show that for workers starting in a white collar occupation who are observed switching a 10% increase in their probability of initially choosing white collar increases the time by on average 0.14 years. For workers starting in blue collar an increase in their probability of choosing white collar decreases the average time to their first switch by a the same amount. In the second specification where the probabilities are constructed using parental occupation variables the point estimates are of the same sign and slightly smaller, but the results are no longer significant as the sample size drops considerably.

Taken together, these results suggest that initial worker characteristics are informative both about worker's initial occupational decisions, but also have lasting consequences in terms of how likely it is that an individual would switch occupations. Workers that are more likely to choose one occupation are less likely to ever switch occupations and conditional on switching take longer to change.

2.2 Occupational and Job Switches

Next, this paper documents the observed patterns of occupational and job switches for high school graduates. [Figure 1](#) shows histograms for the number of years that individuals are observed making job and occupational switches between the ages of 19 and 28. The left panel shows the number of years that each individual has at least one job switch. The right panel shows the number of years where each individual is observed to switch between

blue and white collar occupations. [Figure 1](#) reaffirms the known fact that individuals switch jobs much more often than they switch occupations. The mean for the number of years (out of 10 possible) that include a job switch is 4.6 while the mean number of years with an occupational switch is 1.3. Only about 2% of individuals do not switch jobs during the 10 years while about 42% do not switch occupations. 19% of workers switch occupations exactly one time. However, despite only having two broad occupational categories, a large portion of workers switches more than once. Another 19% switch twice and the final 20% of the population switches more than twice. This means that a sizable portion of the population switches back and forth between blue and white collar jobs during their first 10 years of employment.

To further understand the workers who switch back and forth, the timing of occupational switches is examined. For each worker, the number of years between occupational switches is recorded for each switch. [Table 5](#) reports the average time for workers to make their first and second switches conditional on observing them switch at least twice during the ten years. Results are reported by their choice of initial occupation. The table also reports the number of observations in each group and the t-statistic from a test of the time to first switch being different than the time to second switch. In contrast with the well known decline in occupational transitions with age, the life-cycle data show that the average time to second switch is on average lower than the time to first switch. One might worry that measurement issues bias these results, however the findings are robust to alternate measurement schemes. The sample is truncated to a 10 year period. This truncation leads to biases in both directions. First, since the first occupation is recorded at age 19 workers could have started in their initial occupation earlier. This effect will tend to decrease the time to first switch variable understating the observed differences⁶. Second, since the data is truncated at 10 years, occupational transitions that occur later in life are omitted. If later first or second switches are missed this will cause the average switch times to increase, but it is unclear if

⁶One might also worry that this approach misses occupational switches before the age of 19 and hence biases the results in a more complicated way. These effects however do not change the overall patterns as they show up in the data without selecting the sample to provide uniform observations across individuals.

the effect would be larger for the time to first or second switch as it depends on the fraction of people who are done switching occupations. The same patterns continue to hold even when the sample is further restricted to those who have at least three switches during the first 10 years.

One interpretation of this result is that it provides support for a learning model where workers are initially unsure of which type of occupation that they are suited for. Based on initial information they make an initial occupational choice and learn over time. This framework can generate a threshold belief about the type where workers will chose occupations based on being above or below the threshold. Such a model will be developed and analyzed in the remainder of the paper.

3 Model

This section develops a life-cycle learning model of occupational choices. The model will extend the framework of [Johnson \(1978\)](#) by extending the number of periods to understand life cycle patterns of mobility. To characterize occupational decisions of workers with heterogeneous initial beliefs it will be necessary to relax the primary assumptions in [Johnson \(1978\)](#) that workers are risk neutral and face a decision between two occupations that have the same mean earnings. The goal of the exercise is to see if a canonical model of learning and occupational choices is capable of matching the observed life cycle patterns of occupational choice.

3.1 Preferences and Technology

Consider an individual who lives for Y periods with period utility function:

$$u(c) = \frac{e^{-\gamma c} - 1}{-\gamma}, \gamma > 0$$

The agent can be one of two types. Let $\mu \in \{0, 1\}$ denote the type of the agent. The

agents type is unknown. The agent's initial belief about her type is given by:

$$p_0 = Pr[\mu = 1|\mathcal{X}_0]$$

\mathcal{X}_0 is the agents information set when they begin work and can include their educational history, race, sex, parental education, parental income, etc. For simplicity, we will let p_{-1} be the probability that each agent is of type 1 and update their beliefs with initial information to match observed choices into blue and white collar occupations.

Agents can choose to work in one of two sectors, $i \in \{B, W\}$. Where $i = W$ is better suited for type 1 and $i = B$ is better for type 0 workers. The sectors can be thought of as “white” and “blue” collar respectively.

In occupation i and idiosyncratic state s^7 , output for a worker of type μ is given by:

$$x_{is}^\mu = \bar{x}_{is}^\mu + \varepsilon_i$$

Where \bar{x}_{is}^μ is the average output for a worker of type μ in occupation i and state s and $\varepsilon_i \sim N(0, \sigma_i^2)$ is independently and identically distributed occupation specific noise on the wage process that does not depend on the state. Denote the CDF of output of a worker in sector i and state s as:

$$G_{is}x|\mu) \sim N(\bar{x}_{is}^\mu, \sigma_i^2)$$

Wages are a function of both the worker's sector i , the state s , and the worker's true type μ .

Finally, if a worker switches occupations she must pay a fixed cost in units of consumption u . This cost can be interpreted as costs of search or lost wages involved in switching occupations. A worker's assets evolve according to:

$$a' = (1 + r)a + w_{is}^\mu - c - \mathbb{I}_{switch}u$$

⁷Here we include general notation for individual level sector specific wage shocks. These can be thought of reflecting local labor market conditions or compensating differentials for individual occupational choice.

where a' is next periods assets and w_{is}^μ are the wages received by the worker of type μ in occupation i when the state is s .

3.2 Wages

To close the model, worker's are paid their output in each period. This implies that wages are given by:

$$w_{is}^\mu = x_{is}^\mu$$

for each sector and worker type. Under this assumption firms will earn zero profits as workers are paid the entire output in each period. This occurs in a competitive equilibrium where there is free entry of firms.

3.3 Learning

Based on their observed output, workers update their beliefs each period. Given the normality of output noise, for any belief, p , the expected distribution of output for a worker in occupation i in state s is given by:

$$\psi_{is}(x, p) = p \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \bar{x}_{is}^1}{\sigma_i} \right)^2} + (1 - p) \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \bar{x}_{is}^0}{\sigma_i} \right)^2}$$

With probability p output is drawn from a normal distribution with mean \bar{x}_{is}^1 and variance σ_i , while with probability $1 - p$ it is drawn from a normal with mean \bar{x}_{is}^0 and the same variance.

Using this known distribution of output, the worker observes her output and uses it to update her belief about the probability that she is a worker of type 1 using Bayes' rule. While working in sector i in state s , given any current belief, p , and observed output for a

given period, x , the updated belief, p' , is formed by conducting a probability ratio test:

$$f_{is}(p, x) \equiv p' = \frac{pe^{-\frac{1}{2}\left(\frac{x-\bar{x}_{is}^1}{\sigma_i}\right)^2}}{pe^{-\frac{1}{2}\left(\frac{x-\bar{x}_{is}^1}{\sigma_i}\right)^2} + (1-p)e^{-\frac{1}{2}\left(\frac{x-\bar{x}_{is}^0}{\sigma_i}\right)^2}}$$

Here the numerator is proportional to the probability of observing output x for a high ability worker in sector i and state s and the denominator is proportional to the total probability of observing output x .

With this updating function, define the inverse function $f^{-1}(p'; p)$ to be x required to have posterior p' given prior p . Then the p.d.f. of expected beliefs next period is given by:

$$\psi(f^{-1}(p'|p)|p) \left| \frac{df^{-1}(p'|p)}{dp'} \right|$$

3.4 Value Functions

At time 0, an agents initial information set is given her asset level and initial belief about her ability: $\mathcal{F}_0 = \{a_0, p_0\}$. Denote the value function for an agent in sector i with assets, a , output, x , state, s , updated belief, p , and age, y , as $V_i(a, x, s, p, y)$. Then at time zero the agent chooses to enter either the blue or white collar occupation before realizing the initial state or wage draw. She solves:

$$\max \left\{ \mathbb{E}_s \mathbb{E}_{x|p_0} V_B(a, x, s, p, 0), \mathbb{E}_s \mathbb{E}_{x|p_0} V_W(a, x, s, p, 0) \right\}$$

Each period the agent receives a wage draw, updates her beliefs about her probability, chooses her optimal consumption and asset holdings and then decides on a career for next period based on her updated belief about her type. After time 0, the value function can be written as:

$$V_i(a, x, s, p, y) = \max_{a'} \frac{e^{-\gamma((1+r)a+x-a'-\mathbb{I}_{switch}u)} - 1}{-\gamma} + \frac{1}{1+r} \tilde{V}(a', s, p, y)$$

Where H_i is the distribution of wage draws given current belief p .

$$H_i(x, p) = pG_i(x|\mu = 1) + (1 - p)G_i(x|\mu = 0)$$

The first term is the agent's period utility function with consumption substituted out using the law of motion for assets and \tilde{V} is the agent's continuation value. Since the agent updates her beliefs about her type based on her wage realization in the current period and can choose to switch occupations based on this realization, when $y < Y$, \tilde{V} is given by:

$$\tilde{V}(a', s, p, y) = \max \left\{ \mathbb{E}_{s'|s} \int V_i(a', x, s', p', y + 1) H(dx, p), \mathbb{E}_{s'|s} \int V_{-i}(a', x, s', p', y + 1) H(dx, p) \right\}$$

When $y = Y$, \tilde{V} is given by:

$$\tilde{V}(a', s, p, Y + 1) = \frac{1 + r}{-\gamma r} e^{-\gamma(ra' + R)} + \frac{1 + r}{\gamma r}$$

Where R is the retirement benefit earned by the worker at the end of her career.

Proposition 1 *The value function can be written as:*

$$V_i(a, x, s, p, y) = \frac{1 + r}{-r\gamma} e^{-\gamma(ra + v_i(x, s, p, y))} + \frac{1 + r}{r\gamma}$$

Where when $y < Y$, $v_i(x, s, p, y)$ solves the recursive equation given by:

$$v_i(x, s, p, y) \equiv \frac{\tilde{v}_i(x, s, p, y) + rx - r\mathbb{I}_{switch}u}{1 + r}$$

and

$$\begin{aligned} \tilde{v}_i(x, s, p, y) = & -\frac{1}{\gamma} \ln \left[-\max \left\{ \mathbb{E}_{s'|s} \int -e^{-\gamma v_i(x, s', p', y+1)} H(dx, p), \right. \right. \\ & \left. \left. \mathbb{E}_{s'|s} \int -e^{-\gamma v_{-i}(x, s', p', y+1)} H(dx, p) e^{\gamma \frac{r}{1+r} u} \right\} \right] \end{aligned} \quad (1)$$

When $y = Y$, $v_i(x, s, p, y)$ is given by:

$$v_i(p, y) = \frac{1+r}{-\gamma r} e^{-\gamma \left(\frac{r}{1+r} (x-R) + R \right)}$$

Proof. See Appendix A. ■

Proposition 1 shows that a worker's optimal occupational decisions are independent of their current asset holdings. This allows us to compute a simpler model by eliminating assets from the state space.

Next, **Proposition 2** solves for the value function as $\gamma \rightarrow 0$. This is currently done without the shocks and needs to be revised with the additional notation.

Proposition 2

$$\lim_{\gamma \downarrow 0} V_i(a, p, y) = \begin{cases} \frac{1+r}{r} (ra + v_i(p, y)) & \text{if } y < Y; \\ \frac{1+r}{r} \left(ra + \frac{r}{1+r} (w_i - R) + R \right) & \text{if } y = Y. \end{cases}$$

where

$$v_i(p, y) = \frac{\lim_{\gamma \downarrow 0} \tilde{v}_i(p, y) + r w_i}{1+r}$$

and

$$\lim_{\gamma \downarrow 0} \tilde{v}_i(p, y) = \int \max \{v_i(p', y+1), v_{-i}(p', y+1)\} H_i(dx, p)$$

Proof. See Appendix B. ■

3.5 Optimal Policy

For every age y the optimal policy is characterized by a collection of thresholds $\{\bar{p}_{is}(y)\}_{i \in \{B, W\}}$ independent of current wealth a . The policy is such that an individual currently working

in the blue sector moves to the high sector if $p \geq \bar{p}_{Bs}(y)$. Similarly, individual currently working in the high sector moves to the low sector if $p \leq \bar{p}_{Ws}(y)$. In the absence of switching costs, when $u = 0$, $\bar{p}_{Bs}(y) = \bar{p}_{Ws}(y) = \bar{p}_s(y)$

Next, **Proposition 3** analyzes the expected time for workers to switch occupations. The basic result is that for two workers in the same occupation, the same state, and with the same age, the worker with the belief furthest away from the threshold will on average take longer to switch occupations. Let $\hat{\tau}$ denote the years left prior to leaving current occupation i .

Proposition 3 *Let $p^d(y)$ and $p^j(y)$ denote the beliefs at age y of individuals d and j . If $p^d(y) > p^j(y)$ then $E\{\hat{\tau}|p^d(y), i = B\} \leq E\{\hat{\tau}|p^j(y), i = B\}$ and $E\{\hat{\tau}|p^d(y), i = W\} \geq E\{\hat{\tau}|p^j(y), i = W\}$.*

Proof. See Appendix C. ■

4 Parameterization

As an initial parameterization we solve the model with no cost of switching, $u = 0$ and no idiosyncratic shocks to see how much switching the baseline model can generate. We will then add shocks in the form of a wage shifter in the white collar sector to investigate if these shocks can allow the model to match the amount of switching found in the data.

We use data from the National Longitudinal Survey of Youth 1979 (NLSY79) to set parameters of the model and test model predictions. With the choice of the following parameters the model can be computed: the number of periods in a worker's career, Y , the interest rate, r , the amount of risk aversion, γ , expected output for each worker type in each sector, \bar{x}_W^1 , \bar{x}_W^0 , \bar{x}_B^1 , and \bar{x}_B^0 , the standard deviation of output in each sector, σ_W and σ_B , the retirement benefit, R , and the fixed cost of switching occupations, u .

In the baseline case let $u = 0$. We normalize the period to be one month so that $r = 0.00327$. The model is solved for the first 10 years of the worker's career so $Y = 120$.

Let $R = 10$ as this parameter is irrelevant for individual's occupational choices. Finally, in the baseline specification $\gamma = 1$.

To set wage parameters we need to make assumptions about the observed wage data from NLSY79. Observed wages in blue and white collar occupations in the model come from a mixture distribution as there are both type 0 and 1 workers in each occupation at each date. To set the initial wage parameters we can assume that individual occupational choices at age 28 correspond with their actual type then set each wage variable to be equal to the median of the group relative wage at age 19. That is $\bar{x}_B^0 = 0.9225$ is the median relative wage of workers in blue collar at age 19 conditional on being blue collar workers at age 28. $\bar{x}_B^1 = 0.8785$ is the group median for blue collar workers at age 19 who are in white collar at age 28. Finally, for workers in white collar $\bar{x}_W^0 = 0.9030$, $\bar{x}_W^1 = 0.9080$.

Next, using the share of workers in each sector at age 19, we can use the observed standard deviations in each sector to pick σ_B and σ_W . First, using the parameterized wage values we can solve for the mean wage in each occupation using the mixture distribution formula:

$$\mu_i = w_i^1 \mu_i^1 + w_i^0 \mu_i^0$$

Where w_i^1 and w_i^0 are the shares of type 0 and 1 workers in each sector respectively. With this we can use the mixture distribution formula for variance to compute the standard deviation in each sector σ_i :

$$\bar{\sigma}_i^2 = w_i^1(\mu_i^1 + \sigma_i^2) + w_i^0(\mu_i^0 + \sigma_i^2) - \mu_i^2$$

With this method we get $\sigma_W = 0.177$ and $\sigma_B = 0.274$.

Finally, we need to parameterize the initial distribution of beliefs. We assume that before choosing an initial occupation the worker observes a signal equivalent to α periods of output in sector L. Updating p_{-1} , this generates a non-degenerate distribution of initial beliefs p_0 . We jointly choose $\alpha = 30$ and $p_{-1} = 0.615$ to match the standard deviation of the fitted probabilities of 0.19 and the initial proportion of 36.4% of workers in white collar.

Figure 2 shows the optimal decision rule for a worker by age, $\bar{p}(y)$, under the baseline

calibration and a histogram of worker's initial beliefs p_0 .

In the second parameterization we consider a simple version of idiosyncratic shocks. Assume that wages in the white collar sector can be in one of two possible states. In the low state, 0, wages are multiplied by $(1 - \Delta)$, and in the high state, 1, they are multiplied by $(1 + \Delta)$. Let the state remain the same in the next period with probability ρ . We are going to consider shocks with $\rho = 0.7$ and $\Delta = 0.005$. This amounts to one percent shifts in wages in the white collar sector with moderate persistence. Finally, let the initial state be 1 with probability 0.5 and the initial distribution is calibrated to $\alpha = 30$ and $p_{-1} = 0.56$ to match the same targets.

Figure 3 shows the decision rules for a worker in each state in the model with the shocks and the initial distribution of wages. The figure shows that the threshold in the high state, $\bar{p}_1(y)$, is lower as being in white collar is more attractive. Recall that we assume that the shocks are idiosyncratic across workers, so each worker is being hit with a different history of shocks that move their age specific thresholds around.

5 Results

This section reports the simulated result for the model under the baseline parameterization and the model with shocks.

5.1 Baseline Parameterization

For the baseline parameterization, we want to evaluate if the model can generate the quantity of occupational switches found in the data. **Figure 4** shows the histogram of the number of switches for workers simulated from the model. The model is unable to generate a sizable amount of switches. Nearly 70% of workers do not switch occupations at all while almost 30% switch exactly once. Only about 2% of workers switch more than once compared to compared to almost 40% in the data.

Another way to understand the amount of switching generated is to plot the hazard rate

of switching out of each occupation for workers of each age. [Figure 5](#) plots the switching rate for white and blue collar workers and compares the results to the data. Confirming the results from the histogram, the annual switching rates for both blue and white collar workers are lower in the model than in the data.

Finally, we can compute the average time of the first and second switch in the model. The simulation generates an average time to first switch of 1.93 years and an average time to the second switch of 1.68 years. This compares to 2.88 and 1.90 years in the data.

5.2 Shocks

As the baseline calibration is unable to generate the large portion of workers that frequently switch occupations that are observed in the data, this section explores the model with a wage shifter in the white collar occupation to see if it is able to generate more switches. [Figure 6](#) shows the simulated histogram of the number of years of occupational switches from the model. Over 50% of workers still do not switch at all and just over 20% switch exactly once. However, 24.4% of workers now switch more than once, which is much closer to the 40% observed in the data.

Next, we plot the annual switching rates for blue and white collar workers against the data. [Figure 7](#) shows that with shocks the switching rates for blue collar workers closely approximates the data, while the switching rates for white collar workers start at the same level then decline rapidly. This is a result of white collar workers not learning very much in the particular parameterization that is being used. This issue is discussed further in the next section.

Finally, the average time to the first and second switches in the model with shocks are 2.24 and 1.90 years respectively. This is even closer to the observed timing in the data.

6 Discussion

This paper documents new life cycle features of occupational choices: worker's initial information is informative both about initial occupational decisions and future decisions to switch occupations, there is a sizable portion of workers that switch occupations frequently, and the average time to a worker's first occupational switch is longer than the average time to the second switch. The first and third can be captured by a broad class of models where workers' optimal decision rule is governed by a switching threshold. This paper writes down a life cycle version of such a model to evaluate if it can generate as many switches as are observed in the data. The baseline parameterization indicates that the model does not generate a large amount of occupational switching. The addition of idiosyncratic wage shocks can help the model to generate the amount of switching observed in the data.

To dig a bit deeper into the reasons for switching, from the perspective of the model switching can occur either because a worker's belief changes such that it crosses a threshold or if the threshold shifts causing the mass of workers to change occupations. Even when switching costs are low (zero), the model does not generate a large amount of switching from workers' beliefs repeatedly crossing the threshold because worker's beliefs on average converge to their true beliefs of 0 or 1. However, shocks can be a useful feature to generate switching as it moves the threshold around. These results could also be viewed as weak evidence that the costs of switching occupations are small.

Finally, the current parameterization is preliminary in a number of dimensions. First, it should be noted that the wage parameters are picked to match the cross section of wages for workers at age 19. In particular, this targeting of the cross section wage distribution determines the speed of learning in the model while not targeting anything about the nature of learning. Additionally, there is no clear unit of time implied by this calibration, so the model has workers learn an amount based on the wage process in each period and the time period of one month has been chosen arbitrarily. A second feature is that workers do not learn very much in the white collar sector as the wages are nearly identical across types. [Figure 8](#)

shows the distribution of beliefs at the end of the simulation, when $Y = 120$. Notice that worker's in blue collar have learned much more than those in white collar as many of their beliefs have nearly converged to 0 while the right hand side of the initial distribution is similar to the initial distribution. Future calibrations will need to more carefully document the amount of learning in each sector to provide a better parameterization that can be used to better understand switching behavior.

References

- Antonovics, Kate and Limor Golan**, “Experimentation and Job Choice,” Discussion Papers, UC San Diego 2010.
- Gathmann, Christina and Uta Schönberg**, “How General Is Human Capital? A Task-Based Approach,” *Journal of Labor Economics*, 01 2010, 28 (1), 1–49.
- Groes, Fane, Philipp Kircher, and Iouri Manovskii**, “The U-Shapes of Occupational Mobility,” Discussion Papers, University of Pennsylvania 2010.
- Hoffmann, Florian**, “An Empirical Model of Life-Cycle Earnings and Mobility Dynamics,” Mimeo, University of British Columbia 2010.
- James, Jonathan**, “Ability Matching and Occupational Choice,” Working Paper 1125, Federal Reserve Bank of Cleveland 2011.
- Johnson, William R.**, “A Theory of Job Shopping,” *The Quarterly Journal of Economics*, May 1978, 92 (2), 261–78.
- Jovanovic, Boyan and Yaw Nyarko**, “Stepping-stone mobility,” *Carnegie-Rochester Conference Series on Public Policy*, June 1997, 46 (1), 289–325.
- Kambourov, Gueorgui and Iouri Manovskii**, “Rising Occupational And Industry Mobility In The United States: 1968-97,” *International Economic Review*, 02 2008, 49 (1), 41–79.
- **and** –, “Occupational Mobility and Wage Inequality,” *Review of Economic Studies*, 04 2009, 76 (2), 731–759.
- **and** –, “Occupational Specificity Of Human Capital,” *International Economic Review*, 02 2009, 50 (1), 63–115.
- Keane, Michael P and Kenneth I Wolpin**, “The Career Decisions of Young Men,” *Journal of Political Economy*, June 1997, 105 (3), 473–522.
- McCall, Brian P.**, “Occupational Matching: A Test of Sorts,” *Journal of Political Economy*, February 1990, 98 (1), 45–69.
- Miller, Robert A.**, “Job Matching and Occupational Choice,” *Journal of Political Economy*, December 1984, 92 (6), 1086–120.
- Neal, Derek**, “Industry-Specific Human Capital: Evidence from Displaced Workers,” *Journal of Labor Economics*, October 1995, 13 (4), 653–77.
- , “The Complexity of Job Mobility among Young Men,” *Journal of Labor Economics*, April 1999, 17 (2), 237–61.

- Papageorgiou, Theodore**, “Learning Your Comparative Advantage,” Mimeo, Penn State University 2010.
- Parent, Daniel**, “Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics,” *Journal of Labor Economics*, April 2000, *18* (2), 306–23.
- Poletaev, Maxim and Chris Robinson**, “Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000,” *Journal of Labor Economics*, 07 2008, *26* (3), 387–420.
- Sullivan, Paul**, “A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search,” *International Economic Review*, February 2010, *51* (1), 289–317.
- Topel, Robert H and Michael P Ward**, “Job Mobility and the Careers of Young Men,” *The Quarterly Journal of Economics*, May 1992, *107* (2), 439–79.

Appendix

A Proof of Proposition 1

Taking the first order condition from the worker's period asset choice problem gives:

$$e^{-\gamma((1+r)a-a'+x-\mathbb{I}_{switch}u)} = \frac{1}{1+r} \frac{d\tilde{V}_i(a', s, p, y)}{da'}$$

Substituting into the value function implies:

$$V_i(a, x, s, p, y) = -\frac{1}{\gamma(1+r)} \frac{d\tilde{V}_i(a', s, p, y)}{da'} + \frac{1}{\gamma} + \frac{1}{1+r} \tilde{V}(a', s, p, y)$$

We assume that the continuation value takes the form $\tilde{V}_i(a', s, p, y) = \frac{1+r}{-r\gamma} e^{-\gamma(ra'+\tilde{v}_i(x,s,p,y))} + \frac{1+r}{r\gamma}$. We proceed in two cases. First, when $y < Y$ we have:

$$\begin{aligned} V_i(a, x, s, p, y) &= -\frac{1}{\gamma(1+r)} \frac{d\left(\frac{1+r}{-r\gamma} e^{-\gamma(ra'+\tilde{v}_i(x,s,p,y))} + \frac{1+r}{r\gamma}\right)}{da'} \\ &\quad + \frac{1}{\gamma} + \frac{1}{1+r} \frac{1+r}{-r\gamma} e^{-\gamma(ra'+\tilde{v}_i(x,s,p,y))} + \frac{1}{r\gamma} \\ &= \frac{1+r}{-r\gamma} e^{-\gamma(ra'+\tilde{v}_i(x,s,p,y))} + \frac{1+r}{r\gamma} \end{aligned}$$

Note that the first order condition reduces to,

$$a' = a + \frac{x}{1+r} - \frac{\tilde{v}_i(x, s, p, y)}{1+r} - \frac{\mathbb{I}_{switch}u}{1+r}$$

which provides the savings rule. Substituting into $V_i(a, x, s, p, y)$ yields,

$$V_i(a, x, s, p, y) = \frac{1+r}{-r\gamma} e^{-\gamma(ra+v_i(x,s,p,y))} + \frac{1+r}{r\gamma}$$

with

$$v_i(p, y) \equiv \frac{\tilde{v}_i(x, s, p, y) + rx - \mathbb{I}_{switch}u}{1+r}$$

Finally,

$$\begin{aligned}
\frac{1+r}{-r\gamma} e^{-\gamma(ra' + \tilde{v}_i(x,s,p,y))} + \frac{1+r}{r\gamma} &= \tilde{V}_i(a', s, p, y) \\
\frac{1+r}{-r\gamma} e^{-\gamma(ra' + \tilde{v}_i(x,s,p,y))} + \frac{1+r}{r\gamma} &= \max \left\{ \mathbb{E}_{s'|s} \int V_i(\cdot) H(dx, p), \mathbb{E}_{s'|s} \int V_{-i}(\cdot) H(dx, p) \right\} \\
-e^{-\gamma(ra' + \tilde{v}_i(x,s,p,y))} &= \max \left\{ \mathbb{E}_{s'|s} \int -e^{-\gamma(ra' + v_i(x,s',p',y+1))} H(dx, p), \right. \\
&\quad \left. \mathbb{E}_{s'|s} \int -e^{-\gamma(ra' + v_{-i}(x,s',p',y+1))} H(dx, p) \right\}
\end{aligned}$$

solving for $\tilde{v}_i(x, s, p, y)$,

$$\begin{aligned}
\tilde{v}_i(x, s, p, y) &= -\frac{1}{\gamma} \ln \left[-\max \left\{ \mathbb{E}_{s'|s} \int -e^{-\gamma v_i(x,s',p',y+1)} H(dx, p), \right. \right. \\
&\quad \left. \left. \mathbb{E}_{s'|s} \int -e^{-\gamma v_{-i}(x,s',p',y+1)} H(dx, p) e^{\gamma \frac{r}{1+r} u} \right\} \right]
\end{aligned}$$

When $y = Y$ the value function is given by:

$$V_i(a, x, s, p, y) = \max_{a'} \frac{e^{-\gamma((1+r)a + x - a' - \mathbb{I}_{switch}u)} - 1}{-\gamma} + \frac{1}{-\gamma r} e^{-\gamma(ra' + R)} + \frac{1}{\gamma r}$$

The first order condition gives:

$$e^{-\gamma((1+r)a + x - a' - \mathbb{I}_{switch}u)} = e^{-\gamma(ra' + R)}$$

Solving for a' gives:

$$a' = a + \frac{x - R - \mathbb{I}_{switch}u}{1+r}$$

Substituting back into the value function gives:

$$V_i(a, x, sp, y) = \frac{1+r}{-\gamma r} \left[e^{-\gamma(ra + \frac{r}{1+r}(x - R - \mathbb{I}_{switch}u) + R)} - 1 \right]$$

This reduces to the result.

B Proof of Proposition 2

Let $b(x, p)$ denote the bayes' rule updating so that $p' = b(x, p)$. By definition, $b(x, p)$ is compact and therefore $e^{-\gamma v_i(b(x,p), y+1)}$ and $e^{-\gamma v_{-i}(b(x,p), y+1)}$ are bounded. Each of these

functions can be approximated by Taylor expansions (these functions are also differentiable),

$$\begin{aligned} e^{-\gamma v_i(b(x,p),y+1)} &\approx 1 - \gamma v_i(b(x,p),y+1) + O(-\gamma v_i(b(x,p),y+1)) \\ e^{-\gamma v_{-i}(b(x,p),y+1)} &\approx 1 - \gamma v_{-i}(b(x,p),y+1) + O(-\gamma v_{-i}(b(x,p),y+1)) \end{aligned}$$

where $O(-\gamma v_i(b(x,p),y+1))$ and $O(-\gamma v_{-i}(b(x,p),y+1))$ are both sequences of terms of order higher than one. Further, each of these functions is a convergent series and thus are bounded. Rewrite [equation \(1\)](#) as

$$\tilde{v}_i(p,y) = -\frac{1}{\gamma} \ln \left[\int (1 - \max \{J_i(\gamma, x, p, y+1), J_{-i}(\gamma, x, p, y+1)\}) H_i(dx, p) \right]$$

where

$$\begin{aligned} J_i(\gamma, x, p, y+1) &= \gamma \left(v_i(b(x,p),y+1) - \frac{O(-\gamma v_i(b(x,p),y+1))}{\gamma} \right) \\ J_{-i}(\gamma, x, p, y+1) &= \gamma \left(v_{-i}(b(x,p),y+1) - \frac{O(-\gamma v_{-i}(b(x,p),y+1))}{\gamma} \right) \end{aligned}$$

The proposition follows by taking the limit when γ approaches 0. L'Hopital is required as $\lim_{\gamma \downarrow 0} \tilde{v}_i(p', y+1, \gamma) = \frac{0}{0}$. The issue is whether the function is differentiable with respect to γ and, in the affirmative case, to characterize its derivative.

Lemma 1

$$\int_x \max \{J_i(\gamma, x, p, y+1), J_{-i}(\gamma, x, p, y+1)\} H_i(dx, p) \quad (2)$$

differentiable with respect to γ .

Proof. Let $x^i(\gamma, p, y+1)$ denote the threshold for the signal such that a worker in sector i wishes to move to sector $-i$. As $J_i(\gamma, x, p, y+1)$ and $J_{-i}(\gamma, x, p, y+1)$ both continuous and differentiable with respect to γ and x , $x^i(\gamma, p, y+1)$ continuous and differentiable with respect to γ .

Let $Q_i(\gamma, p, y+1) \equiv \int \max \{J_i(\gamma, x, p, y+1), J_{-i}(\gamma, x, p, y+1)\} H_i(dx, p)$. [equation \(2\)](#) can be rewritten as

$$\begin{aligned} &\int_0^{x^L(\gamma,p,y+1)} J_L(\gamma, x, p, y+1) H_L(dx, p) + \int_{x^L(\gamma,p,y+1)}^1 J_H(\gamma, x, p, y+1) H_L(dx, p) \quad \text{for workers at sector } L \\ &\int_0^{x^H(\gamma,p,y+1)} J_L(\gamma, x, p, y+1) H_H(dx, p) + \int_{x^H(\gamma,p,y+1)}^1 J_H(\gamma, x, p, y+1) H_H(dx, p) \quad \text{for workers at sector } H \end{aligned}$$

that is differentiable with respect to γ . ■

Now we characterize the derivative of $Q_i(\gamma, p, y + 1)$ with respect to γ .

Lemma 2

$$\frac{dQ_i(\gamma, p, y + 1)}{d\gamma} = \int_x \max \left\{ \frac{\partial J_i(\gamma, x, p, y + 1)}{\partial \gamma}, \frac{\partial J_{-i}(\gamma, x, p, y + 1)}{\partial \gamma} \right\} H_i(dx, p)$$

Proof. Follows by applying Leibniz's rule and by noting that

$$\begin{aligned} J_L(\gamma, x^L, p, y + 1) \frac{\partial x^L(\gamma, p, y + 1)}{\partial \gamma} &= J_H(\gamma, x^L, p, y + 1) \frac{\partial x^L(\gamma, p, y + 1)}{\partial \gamma} \\ J_L(\gamma, x_H, p, y + 1) \frac{\partial x_H(\gamma, p, y + 1)}{\partial \gamma} &= J_H(\gamma, x_H, p, y + 1) \frac{\partial x_H(\gamma, p, y + 1)}{\partial \gamma} \end{aligned}$$

by construction of the thresholds $x^L(\gamma, p, y + 1)$ and $x_H(\gamma, p, y + 1)$. ■

The proposition follows by applying L'Hopital's rule and using the last result.

C Proof of Proposition 3

We prove the statement for workers currently in the low sector. The proof for workers in the high sector follows by symmetry. The proof follows by a sequence of claims.

Claim 1 *Consider two individuals d and j of the same age y working in the low sector and sharing a common ability level μ . If $p^d(y) > p^j(y)$ then $E\{\hat{\tau}|p^d(y), i = L, \mu\} < E\{\hat{\tau}|p^j(y), i = L, \mu\}$. In other words, d is more likely to leave her current occupation sooner.*

Proof. As both individuals have the same age the time to retirement conditional on staying in the current sector is the same. Therefore, to prove the statement we only need to consider the probability of switching occupations.

As d and j are of the same age they face the same profile of thresholds. Also, as they are both of the same type the densities of possible signals that can arrive are the same for both. As the posterior is increasing in the prior p it follows directly that $\Pr[p(y + 1) \geq \bar{p}_L(y + 1) | p(y) < \bar{p}_L(y), p(y) = p^d(y), \mu] > \Pr[p(y + 1) \geq \bar{p}_L(y + 1) | p(y) < \bar{p}_L(y), p(y) = p^j(y), \mu]$. Conditional on no switching occupations, $p^d(y + \tau) > p^j(y + \tau)$ for any $\tau \leq Y - y$ and the proof follows by induction. ■

Figure 1: *Histograms of the number of years with job and occupation switches between the age of 19 and 28. Left panel shows the histogram of the number of years that contain at least one job switch. The right panel shows the number of years where workers switch occupations between blue and white collar.*

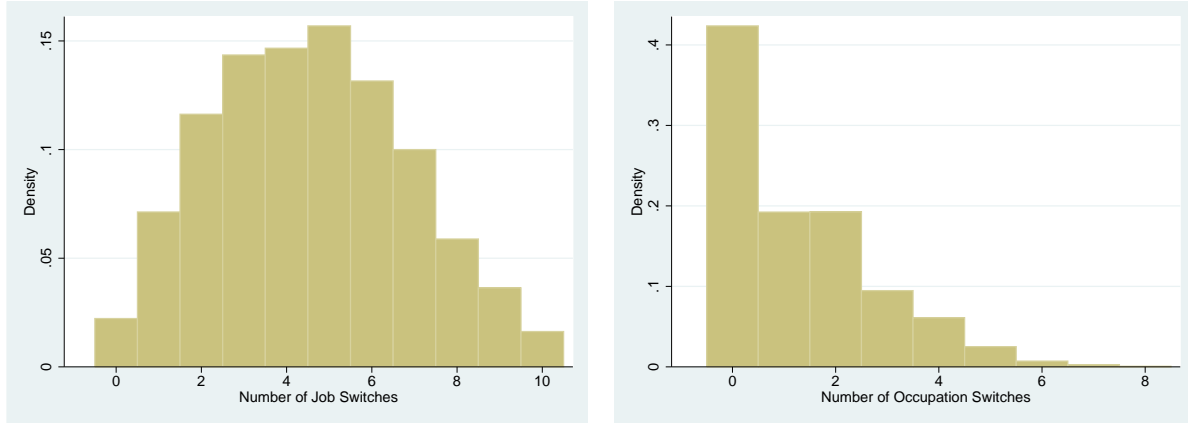


Figure 2: *Baseline $\bar{p}(y)$ and p_0 . The left panel of the figure shows the optimal decision rule for a worker to switch occupations as a function of their age, $\bar{p}(y)$. The right panel shows a histogram of initial beliefs, p_0 , under the baseline calibration.*

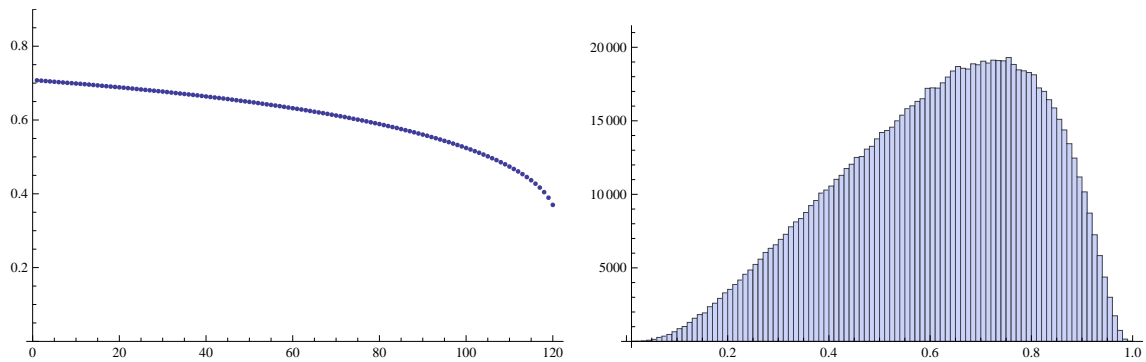


Figure 3: *Optimal decision thresholds with shocks $\bar{p}_0(y)$, $\bar{p}_1(y)$ and initial distribuiton of beliefs p_0 . The left panel of the figure shows the optimal decision rule for a worker to switch occupations as a function of their age and shocks, $\bar{p}_0(y)$ and $\bar{p}_1(y)$. Note that the threshold in the high state, $\bar{p}_1(y)$, is lower as being in white collar is more attractive. The right panel shows a histogram of initial beliefs, p_0 , under the baseline calibration.*

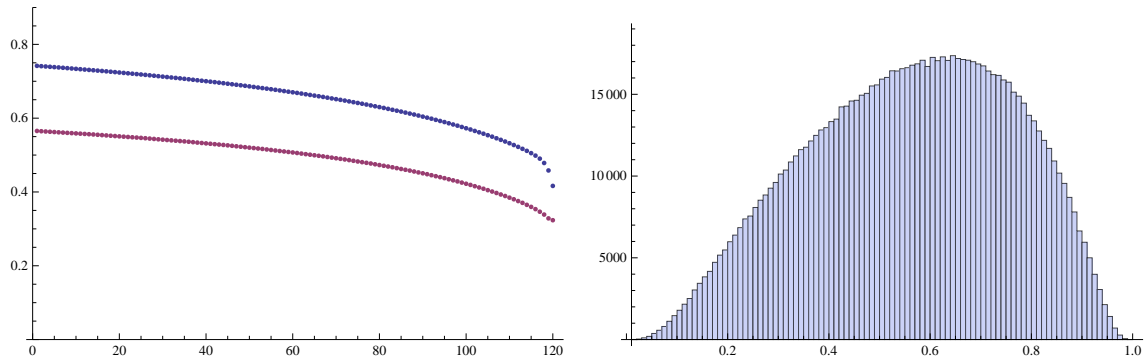


Figure 4: *Distribution of occupational switches in simulated model with baseline parameterization.*

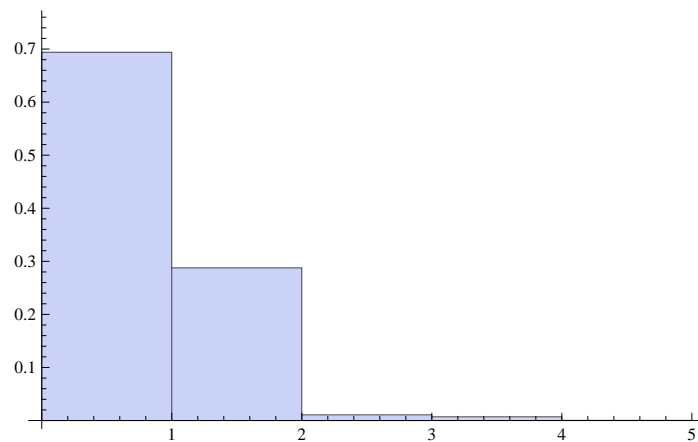


Figure 5: *Switching rates by occupation and age. The left panel shows the simulated switching rates from blue collar to white collar by year. The model is graphed as a blue line while the data are plotted as points. The right panel plots the switching rates from white to blue collar.*

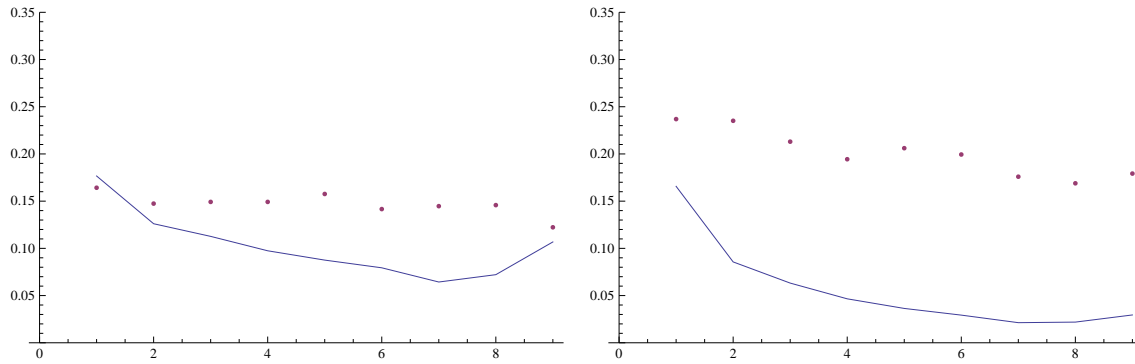


Figure 6: *Distribution of occupational switches in simulated model in parameterization with shocks.*

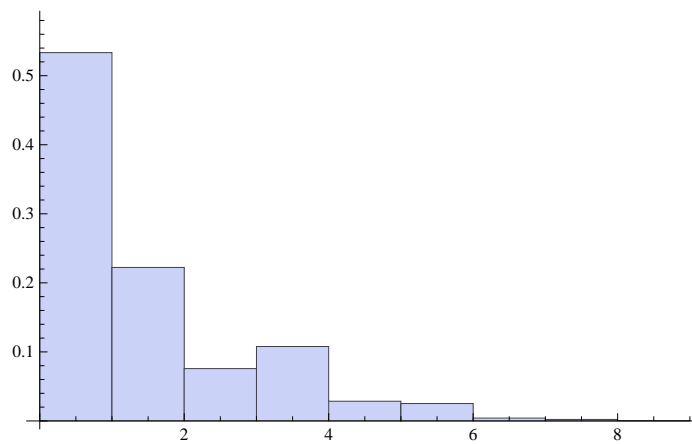


Figure 7: *Switching rates by occupation and age in model with shocks. The left panel shows the simulated switching rates from blue collar to white collar by year. The model is graphed as a blue line while the data are plotted as points. The right panel plots the switching rates from white to blue collar.*

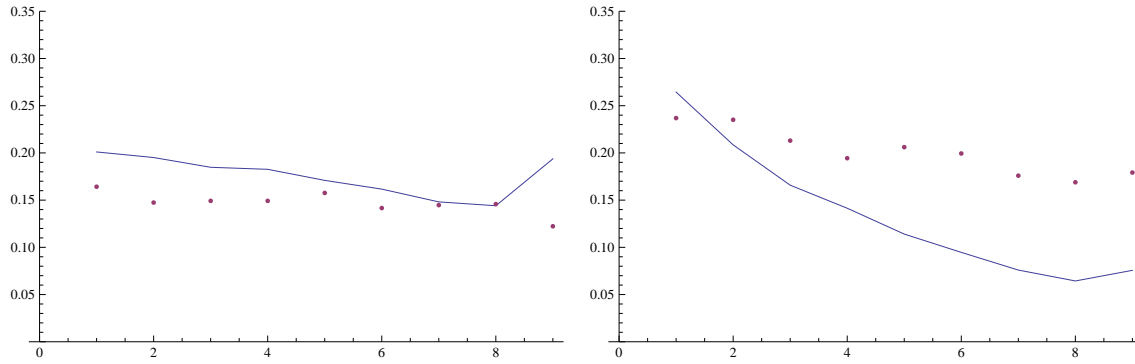


Figure 8: *Distribution of beliefs for baseline parameterization when $Y = 120$.*

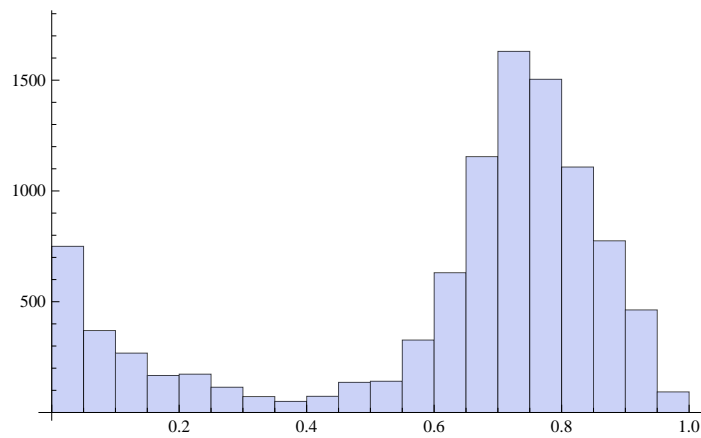


Table 1: *Summary statistics for initial occupational choice for sample with high school degree for highest education received. Results are broken down by initial occupational choice. The table shows the mean of each variable with standard deviation in parentheses. Number of observations for each variable shown in final column.*

Variable	Blue Collar	White Collar	Total	Observations
First Occupation White Collar	0 (0)	1 (0)	0.352 (0.478)	4180
Mother's Years of Schooling	10.51 (2.846)	10.78 (2.817)	10.61 (2.838)	3944
Father's Years of Schooling	10.31 (3.472)	10.89 (3.429)	10.52 (3.468)	3558
Mother's Main Occupation WC	0.355 (0.479)	0.451 (0.498)	0.390 (0.488)	2391
Father's Main Occupation WC	0.244 (0.430)	0.311 (0.463)	0.268 (0.443)	2960
Male	0.628 (0.483)	0.294 (0.456)	0.511 (0.500)	4180
White	0.541 (0.498)	0.532 (0.499)	0.538 (0.499)	4180
Urban	0.748 (0.434)	0.828 (0.377)	0.777 (0.417)	4162
Family Poverty	0.398 (0.489)	0.330 (0.470)	0.374 (0.484)	3998
Class Percentile	0.370 (0.254)	0.481 (0.263)	0.413 (0.263)	2251
AFQT Percentile	0.310 (0.227)	0.378 (0.233)	0.334 (0.231)	4015

Table 2: *Probit regression of choosing white collar as the initial occupation on observable initial characteristics at time of first occupational choice. Marginal effects are reported with standard errors in parentheses. Specification (1) omits parental occupation variables while specification (2) includes them.*

Variables	(1) First Occupation WC	(2) First Occupation WC
Mother's Years of Schooling	-0.00239 (0.00399)	-0.0129* (0.00670)
Father's Years of Schooling	0.0142*** (0.00328)	0.0174*** (0.00548)
Mother's Main Occupation WC		0.0683** (0.0299)
Father's Main Occupation WC		0.0206 (0.0317)
Male	-0.292*** (0.0175)	-0.313*** (0.0255)
White	-0.0808*** (0.0209)	-0.0747** (0.0309)
Urban	0.111*** (0.0203)	0.0953*** (0.0306)
Family Poverty	-0.0385** (0.0195)	-0.0678** (0.0309)
Class Percentile	0.360*** (0.0511)	0.430*** (0.0746)
AFQT Percentile	0.143*** (0.0528)	0.0389 (0.0778)
Observations	3,160	1,490
R^2	0.131	0.144

*** p<0.01, ** p<0.05, * p<0.1

Table 3: *Probit regression of ever switching occupations on the fitted probability of initially choosing a white collar occupation. Results are broken down by initial occupational choice. First two columns present results from the first specification with parental occupation variables dropped. Final two columns show result for specification with probabilities from the probit including parental occupation variables. Marginal effects presented. Standard errors in parentheses.*

	(1)	(1)	(2)	(2)
Variables	White Collar	Blue Collar	White Collar	Blue Collar
Probability	-0.679*** (0.0934)	0.449*** (0.0593)	-0.609*** (0.130)	0.467*** (0.0810)
Observations	934	2,226	457	1,033
*** p<0.01, ** p<0.05, * p<0.1				

Table 4: *Regression of the timing of the first occupational switch on the probability of initially choosing a white collar occupation. Results broken down by initial occupational choice. The first two columns use the fitted probability from the probit regression with the parental occupation variables omitted. The second two columns include these variables in the construction of the fitted probability. Standard errors in parentheses.*

	(1)	(1)	(2)	(2)
Variable	White Collar	Blue Collar	White Collar	Blue Collar
Probability	1.382** (0.548)	-1.392*** (0.376)	1.187 (0.783)	-0.982* (0.506)
Constant	2.373*** (0.260)	3.801*** (0.143)	2.403*** (0.399)	3.545*** (0.201)
Observations	529	1,239	255	594
R-squared	0.012	0.011	0.009	0.006
*** p<0.01, ** p<0.05, * p<0.1				

Table 5: *Average time to first and second occupational switch by initial occupational choice*. Results are conditional on switching at least twice. Mean of each variable with standard deviation in parentheses. Number of observations and t-statistics for a test that the means are different between the two means are reported in each case.

	Blue Collar	White Collar	Total
Time to First Switch	3.179 (2.041)	2.355 (1.798)	2.883 (1.996)
Time to Second Switch	1.732 (1.307)	2.199 (1.760)	1.899 (1.502)
Observations	961	538	1499
t-statistic	16.65 ***	1.27	13.51 ***
*** p<0.01, ** p<0.05, * p<0.1			