Distributional Effects of Ability Learning and Career Choice

Sung Ah Bahk *

March 29, 2023

Abstract

This paper examines occupational choices as a channel between wealth and earnings inequality. Selecting occupations that match workers' skills can lead to higher earnings. However, risk-averse workers might be reluctant to experiment and discover their highest earnings potential due to uncertainty. I construct a dynamic structural model that examines how workers gradually learn about their skills by performing different combinations of tasks. The results show that lower wealth leads to less experimentation in occupational choice, resulting in a greater loss in lifetime earnings. Additionally, the study finds that the payoff for performing cognitive tasks is much greater than that for motor tasks, but the costs of mismatching to a task are also greater for cognitive tasks. Simulation results using the estimated model show that a policy proposal for Baby Bonds can significantly reduce income inequality, particularly for people with low wealth.

Key words: occupational choice, skill uncertainty, learning, wealth inequality

JEL codes: J24, J62, E24

*I deeply thank Robert Moffitt, Yuya Takahashi and Yingyao Hu for invaluable advice and guidance. I am also grateful to seminar and conference participants at California State University, East Bay, Colorado State University, Johns Hopkins University, Ryerson university, University of California, Santa Cruz, Econometric Society Summer School in Dynamic Structural Econometrics, and Econometric Society Asian Meeting. All errors are mine. American University. Email: bahk@american.edu

1 Introduction

Selecting the occupation that best matches one's personal skills allows for optimal job performance, which is crucial not only for the individual but also for aggregate productivity within the larger economy. There are a number of factors that make this important decision difficult. One well-known factor is imperfect information about one's own skills. Workers who lack perfect awareness of the true nature and level of their abilities learn through on-the-job experience, observing how their individual performance is compensated and which types of jobs yield the highest earnings for them. Nevertheless, risk-averse workers may be reluctant to experiment with job changes to identify high-paying positions, leading to inefficiencies due to imperfect information. This inefficiency may disproportionately affect low-wealth workers, exacerbating wealth inequality through earnings inequality.

In this paper, I develop a career choice model that describes optimal job-task selection under skill uncertainty, with a specific focus on the dynamic effects of wealth levels on earnings over the life cycle. The model has some of novel features that are new to existing career choice models. First of all, I use the concept of occupational "distance" to measure degrees of job changes to differentiate between changes in similar occupations and qualitative changes in ones' career. I utilize occupational task information from the Dictionary of Occupational Titles (DOT). Using two dimensional task scores, cognitive and motor, I measure the degree of occupational changes as Euclidean distance.

Secondly, I use wage residuals as a noisy signal of workers' productivity at their current job to reflect the gradual learning that takes place in the labor market[1]. Based on the wage signal, workers update their beliefs about their skills and select their future occupations accordingly. The wage signals are weighted by occupational tasks in the model. That is, occupations that do not require any cognitive tasks do not provide any information about a worker's cognitive skill. Similarly, earnings in occupations that require high cognitive and low motor skills, such as economists, provide more information about a worker's cognitive ability but less about their motor skills, and the task weights account for this.

[1]The idea of workers learning about their personal skill levels is widely discussed in the literature since Jovanovic (1979), empirical evidence also suggests that job mobility decreases with age and tenure (Neal (1999), Antonovics and Golan (2012), Papageorgiou (2014), Gorry et al. (2019)) and with current wage residuals (Arcidiacono et al. (2016) Guvenen et al. (2020)

The model predicts that not only those who receive bad productivity signals but also those who receive good signals are likely to change their occupations. This feature is consistent with the results in Groes et al. (2015) and the reduced form findings presented in the section 7 of this paper. Using the same measures of occupational distance and productivity signals, I provide reduced-form evidence that workers' future occupational choices are highly correlated with wage signals from their current job. Both the directions in changes, such as moving up or down in each task, and the occupational distance are significantly correlated with the wage residuals.

The model parameters are estimated using the National Longitudinal Survey of Youth 1979 (NLSY79) panels. The findings indicate a significant constant relative risk-aversion, meaning that the inefficiencies in occupational choice decrease with workers' level of wealth. Wealth has an important role because the volatility in future income due to unknown individual ability involves different levels of utility costs depending on the wealth. Furthermore, I find that the payoffs for performing cognitive tasks are much larger compared to an equivalent level of motor task. However, the penalty for overshooting a cognitive task – that is, of choosing a cognitive task that is too high for one's true ability – is much larger than it is for motor tasks. The results also demonstrate that workers have higher uncertainty in motor skills at the beginning, but it resolves faster over the life cycle compared to cognitive skills.

Using the estimated model, I simulate a policy proposal for Baby Bonds, which involves a federally seeded trust fund for every U.S. newborn. The simulation results show that supporting young adults at the beginning of their entry into the labor market can have a significant and long-term effect on reducing income inequality. This proposal provides figurative 'insurance' to workers during the process of discovering their comparative advantages within the labor market. The model's results suggest that this policy will be particularly beneficial for people with low initial wealth. With an annual support of \$1,000 for 18 years, the income ratio between the bottom 10% and top 10% of the initial wealth group will increase by 1.01 percentage points. If the annual support is increased to \$2,000, the income ratio will increase by 5.36 percentage points.

2 Relation to the Literature

This paper builds upon two strands of existing literature. The first is the literature on learning and labor market transitions. The idea of skill uncertainty and learning was first proposed in the classic matching model by Jovanovic in 1979. In his model, both workers and employers face uncertainty about workers' skills and it is only after they are matched and begin working is their true productivity revealed. They subsequently decide whether to remain matched or to split up. There has been other empirical work on learning and labor market transition although most of the learning literature has focused on workers' job-specific (Jovanovic (1979), Gorry et al. (2019)), occupation-specific abilities (Kambourov and Manovskii (2009), Pavan (2011), Antonovics and Golan (2012), Papageorgiou (2014)). For example, the process of learning about skills within blue collar versus white collar jobs, or professional occupations versus non-professional occupations, has been widely studied. These studies found that nearly 20% of workers change their occupation every year, and that subsequent wage gains are as large as a third of early-career wage growth. (Kambourov and Manovskii (2009), Topel and Ward (1992)).

A growing body of literature, since Autor et al. (2003), has considered task-specific approaches (Yamaguchi (2012), Sanders (2014), Autor and Dorn (2013)) instead of job or occupation-specific abilities using task information in DOT and its successor, the Occupational Information Network (O*NET). In this literature, skills are assumed to be task specific and to reflect the daily responsibilities of workers. For example, Yamaguchi (2012) and Sanders (2014) consider skills in two continuous dimensions, cognitive and motor. Compared to the job-specific or occupation-specific skills, the task-specific approach has many attractive characteristics for studying occupational mobility, because it allows comparisons among different occupations in terms of skill levels that are required to perform a job. In addition, it is easier to handle a large number of occupations when they are defined in task-specific skills. This paper builds upon these task-specific extensions of Roy (1951) model, and I introduce skill uncertainty and risk aversion to analyze the role of wealth in career choice and the resulting income inequality.

There have been some studies that analyze occupational decisions as risky choices (King (1974), Saks and Shore (2005)). Most works in this vein focus on risk averse individuals' occupation choices, given some distributional characteristics pertaining to occupation, such as the mean or variance of wages within occupations and test whether workers with low wealth are likely to choose occupations that have low wage variances. Therefore they conclude that workers are likely to choose a certain occupation to another depending on their asset holdings. My paper is distinct from these works, in that it treats risk in career choice as deriving from workers' lack of knowledge of their own abilities, rather than from occupation-specific characteristics, and in that the uncertainty is gradually resolved from work experience.

A recent working paper Hawkins and Mustre-del Rio (2016) investigates the effect of market incompleteness on occupational mobility and finds that low-asset workers are reluctant to switch to high productivity occupations. I complement the findings in Hawkins and Mustre-del Rio (2016) by providing a structural model that features the strong life-cycle pattern of occupational mobility through gradual learning about task-specific ability.

3 Data

Research on occupational mobility often struggles with the question of how best to group occupations into categories. Coarse definitions do not entirely capture the differences between and within occupations. There may exist a huge discrepancy among the occupations which are grouped in the same categories, and there may be some occupations in different categories but share similar characteristics in terms of the job tasks. However, finer distinctions (thus resulting in more categories) are hard to operationalize, because the number of parameters or states increase with the number of occupations. In addition to these difficulties, it is hard to evaluate on a practical level how similar or different two distinctly coded occupations are. Therefore, skill transfer across occupations is often ignored when assessing returns to occupational tenure.

To overcome this problem, I borrow Yamaguchi (2012)'s framework, which defines an occupation as a bundle of tasks along two different dimensions, cognitive and motor complexity. The advantage of using task complexity is that it allows for richer evaluations of occupational mobility; not only can the frequency of movement be analyzed, but also the directionality of movement. For example, does the worker move up or down the scale of task complexity (along the cognitive and/or motor dimension), or does the worker perhaps move laterally - into a new job that requires different kinds of skills? Such a categorization method also allows me to observe how radical an occupational transition is, as we can examine the distance between the new and old task requirements.

The following subsections include explanations for the data constructed in Yamaguchi (2012) and the additional variable, initial wealth in detail.

3.1 Dictionary of Occupational Titles

The Dictionary of Occupational Titles (DOT) contains detailed task information on 12,099 occupations. Each occupation is evaluated with respect to 62 characteristics, such as aptitudes, temperaments, necessary training time, physical demand, and working conditions. Yamaguchi (2012), like many other authors who use the DOT, categorizes these job characteristics into cognitive and motor tasks (Bacolod and Blum (2010), Ingram and Neumann (2006)). Autor et al. (2003) and Autor and Dorn (2013) consider three skill dimensions including abstract, manual, and routine task to analyze the allocation of task between labor and capital due to the technological changes in the labor market.

The DOT variables that Yamaguchi uses to measure cognitive complexity consist of two worker function variables (data and people), three general educational development variables (reasoning, mathematical, and language), three aptitude variables (intelligence, verbal, and numerical), and three adaptability variables (influencing people, accepting responsibility for direction, and dealing with people). The motor complexity measure, meanwhile, comes from 20 physical demand variables, including motor coordination, finger dexterity, manual dexterity, eye-hand-foot coordination, spatial perception, form perception, color discrimination, setting limits, and tolerance or standards. Following Autor et al. (2003), the two measures, cognitive and motor complexity, are converted into percentile scores among the all occupations, taking a value between 0 and 1.

3.2 National Longitudinal Survey of Youth 1979

The National Longitudinal Survey of Youth 1979 (NLSY79) is particularly suitable for this study because it is a long panel data set which contains detailed individual career information and focuses on the young, when most labor market transactions actually occur (Neal (1999)). The survey includes individuals who are between 14 and 21 years old as of January 1, 1979. Occupations in the NLSY79 are coded using a three-digit Census frame, which consists of 503 distinctive categories. Yamaguchi restricts his samples to male workers who make long-term transitions in the labor market during 1979-2000. A long-term transition means working 30 hours per week or more for three consecutive years during the periods. The final data set includes 2,417 workers' career history, 32,774 person-year observations of occupational choices and 31,157 person-year observations of wages. The DOT occupations are aggregated into the three-digit classifications in order to merge with the NLSY79. Worker characteristics such as race, AFQT (Armed Forces Qualification Test) score, and years of education are also obtained from the NLSY79. Excluding non-workers from the sample may bias the estimators in the event some people lose a job because of a bad match and their wage is missing.

In addition to occupation, wage, workers' pre-labor market characteristics variables, I obtained initial asset information from the NLSY79. I focus on money assets such as the savings account of the respondent and his/her spouse. Household assets are recorded after 1985 and once every two years in NLSY79. Therefore, for workers who entered in the labor market between 1979 to 1984, I have information about their wealth level only after they have worked and accumulated assets for some years. For this reason, I construct predicted initial wealth using the information from workers who have records of initial assets at their labor market entry, workers' initial characteristics such as years of educational attainment, AFQT scores, demographics, and their first period labor earnings.

Table 1 reports the mean and standard deviation of all variables and Figure 1 shows the histogram of the predicted initial wealth data used in this paper. The mean of the AFQT score is 49.079 and the standard deviation is 30.1438. The average years of educational attainment is slightly over 13 years (13.2375) and its standard deviation is 2.5353. The percentage of Hispanics in the sample is about 11%, and about 82% of the sample are whites. Average age at the labor market entry is 21.1386 with a standard deviation of 2.9532. Log Initial asset levels are in 2005 real dollars. The mean is 5.57 and the standard deviation is 2.6359. As mentioned (asset levels are only recorded after 1985, once in two years), only a smaller subset of the full sample (503 out of 2417) have data on initial assets. The summary of wage and occupations show hourly wage rates in 2005 real dollars, and cognitive and motor task choice observed yearly from 1979 to 2000. The mean of the hourly wage rate is 17.6904 with a standard deviation of 10.4820.

Table 2 shows more detailed summaries for the panel data. We can see that on average, the hourly wage rate increases continuously from 11.8307 in year 1 to 23.4238 in year 22. The variance of the wage distribution increases in years as well, from 5.7708 in year 1 to 12.3660 in year 22. This is a common finding, earnings profiles spread out over time. The number of observations decreases in year because 'year' indicates the years after labor market entry. A smaller number of data points are observed for longer periods of time.

Table 1 also shows the summary of pooled data on occupational choices. The mean of cognitive task choice for all years is 0.5018 and its standard deviation is 0.2645, and the mean and standard deviation for motor tasks is 0.5291 and 0.2487 respectively. Table 2 describes how they change over the life cycle. For cognitive choices, we observe increasing trends from 0.4124 to 0.5436, and slightly decreasing trends in motor tasks from 0.5313 to 0.5147 for over 22 years. It is rising for the first 6 years, and starts to decline after. The on-the-job skill accumulation (learning-by-doing) process could be very different for different sets of skills, suggesting that a single accumulation rule for skills in multi-dimension will not be able to successfully illustrate skill changes over the life-cycle. Another possible explanation is that motor skills are likely to depreciate as workers get older, and this skill depreciation effect may dominate skill accumulation in the later periods of the life-cycle. The standard deviations stay roughly the same for both tasks.

Table 2 shows that variations in the average task choices are moderate overtime, however, the transitions across different level of tasks, which are hidden in the average trend, occur actively over the life cycle. Table 3 shows that a significant fraction of the workers are moving up and down across task ladders, both in cognitive and motor tasks. The table shows transition probabilities across task quintiles where the first quintile of the cognitive (motor) tasks includes all occupations with cognitive (motor) tasks less than 0.2. And the fifth quintile includes all occupations with high level of cognitive (motor) tasks greater or equal to 0.8. The transition probabilities for cognitive tasks show that 57.63% of the people who were in the first quintile of cognitive tasks stay in the same group next year. And the rest switch to occupations with higher cognitive tasks. The chance of staying in the same cognitive task quintile is higher when workers are in high cognitive jobs, however, even in the highest quintile, about 25% of the people move down to lower group in the next year. The transitions in the motor tasks show that about 60% of the people stay in the same group for all five motor task quintiles, which explains why the average trend in motor tasks is stable over time in Table 2. About 40 % of the workers in each motor task quintile move up or down, changing the intensities of motor tasks in their job.

4 Model

This section develops a dynamic career choice model that accounts for learning and wealth inequality. Occupations are defined over the two-dimensional continuous tasks space. Workers observe their current wage as a productivity signal and update their beliefs about their abilities accordingly. Workers are risk-averse and heterogeneous with respect to initial wealth and skill endowments. All information regarding workers' work history is assumed to be public; therefore employers are assumed to have symmetric information. Finally, the labor market is assumed to be competitive.

Both informational friction and risk aversion are crucial for wealth inequality to have a role in the career choice. In the perfect-information case, for example, a worker knows in which occupation he can be most productive and hence knows which career path results in the highest payoffs. Regardless of his risk preference, then, any worker in a perfect-information scenario would choose the occupation that gives the highest future wage streams, to maximize his lifetime budget. Risk preference, in this case, could affect a decision-maker's consumption and savings behavior but not his career choice; therefore wealth would not play a role in occupational choice.

Meanwhile, a worker who is risk-neutral but does not know her ability perfectly will choose an occupation that has the highest expected wage streams regardless how much risk is involved in that choice. A wage-maximizing career path maximizes the risk-neutral worker's lifetime utility as well, and her occupational choice depends only on her own (imperfect) belief in her ability, but not on the wealth in her hands.

If a worker is both risk-averse and has imperfect information about his own ability, however, that worker might not want to choose the wage-maximizing occupation after all, if the wage-maximizing choice were associated with high risk. Workers may be discouraged from actively engaging in learning. Hence, underinvestment in career choice occurs, and the gap between risk-optimal and wage-maximizing occupations could be wider for workers with low wealth under plausible assumptions such as constant relative risk aversion or credit constraints. In effect, workers with greater wealth are more likely to find occupations with better fit, and therefore wealth inequality can increase further still. Additionally, as time passes by, workers who experiment more, learn more about their ability, and so wage inequality may increase even further.

Lifetime income risk, in this model, derives from skill uncertainty; workers make decisions about consumption, savings, and their next-period occupations, all in response to the partial realization of uncertainty, and they adjust their career paths accordingly in order to maximize lifetime utility.

My model consists of the following elements: 1) a wage function which is determined by the choice of tasks, workers' true skills, and a random transitory shock, and 2) a skill accumulation equation which depends on workers' previous skill level, task choices, and a permanent productivity shock on each skill dimension, and 3) a formal description of learning and belief updating process by Bayes Rule. The following subsections describe this model in detail.

4.1 Wage function

Both workers and employers have imperfect and symmetric information about workers' skills, and the labor market is assumed to be competitive. Therefore, workers are paid by their marginal value product. The marginal value product of a worker with skill $s_t = (s_{ct}, s_{mt}) \in \mathbb{R}^2$ in an occupation with task complexity $x_t = (x_{ct}, x_{mt}) \in (0, 1)^2$ is

$$w_t = \pi(x_t) + q(x_t, s_t) + \epsilon_t, \tag{1}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ is an independent and identically distributed transitory productivity shock. The output price from task x_t is defined as $\pi(x_t)$, and $q(x_t, s_t)$ is a worker's marginal productivity which depends on the task x_t and skill s_t levels. I assume that the marginal productivity of a worker who is endowed with skill s_t in occupation x_t takes the following form:

$$q(x_t, s_t) = (B_2(\alpha s_t - x_t))' x_t, \tag{2}$$

where α is a scalar, and B_2 is 2-dimensional diagonal matrix. Labor productivity is the inner product of excessive skill and task requirement. Note that components in the term $(\alpha s_t - x_t)$ can be negative. If a low skilled worker chooses an occupation that requires much higher task, the low labor productivity will result in low wages. The output price $\pi(x_t)$ is assumed to be linear in the task requirements.

$$\pi(x_t) = B_0 + B_1' x_t, \tag{3}$$

where B_1 is a two dimensional vector. Finally, period t wage w_t can be written as

$$\begin{aligned} w_t &= B_0 + B_1' x_t - x_t' B_2 x_t + g_t \\ g_t &= x_t' B_3 s_t + \epsilon_t \end{aligned} \tag{4}$$

where $B_3 = \alpha B_2$. Wage coefficients B_0, B_1, B_2 and B_3, and the distribution of the transitory shocks ϵ_t are known to workers, but they do not observe their true skills s_t and realization of the shock ϵ_t. Therefore, when w_t is unexpectedly high (or low), workers cannot perfectly pin down whether that is because their true skill s_t is high (low) or they were just lucky (unlucky).

By construction, the variance in the wage distribution within an occupation is larger as the job task intensity x_t is higher. For example, for an imaginary occupation $x_{ct} = x_{mt} = 0$, w_t is simply $B_0 + \epsilon_t$, regardless of workers' skill s_t. However, as the task intensity x_t rises, the wage depends more and more heavily on the workers' true skill levels s_t and the variance in the wage distribution becomes larger.

If B_1, B_2 and B_3 are positive, for any worker with $s_t > 0$, the expected value of the unknown part of the wage g_t increases with respect to the occupational choice x_t. However, the marginal change in the known part of the wage, $B_1 - B_2 x_t$ decreases with respect to x_t, hence B_2, cost of mismatch (overshooting), provides an incentive not to choose high x_t for workers who believe their true skills s_t are low.

4.2 Skill accumulation

Skills accumulate based on the workers' skills in the previous period and the choice of occupations. The following equation shows the process of skill accumulation:

$$\begin{aligned} s_{t+1} &= s_t + A_1' x_t + x_t' A_2 x_t + \eta_t \\ s_0 &= H_0 + H_1 d + \eta_0 \end{aligned} \tag{5}$$

Where $\eta_t \sim N(0, \sigma_\eta^2)$ is a iid shock on the skills which reflect a permanent shock in the productivity. Workers know the values of A_1, A_2, and the distribution of the permanent skill shocks, η_t. Initial skills s_0 depend on the worker's characteristics d before the labor market entry and the unknown iid skill shock $\eta_0 \sim N(0, \sigma_{\eta 0}^2)$. A_0 , A_1, H_0, s_t, η_t, and σ_η are 2 dimensional vectors, and A_2 is 2×2 diagonal matrix.

4.3 Learning

Workers do not exactly know their skill levels. Instead, workers have beliefs about their skills and update the belief from the wage realization each period. Wage depends on the labor productivity

$q(x_t, s_t)$, so it is informative about skills. Workers observe a signal g_t, the sum of the last two terms in the wage equation $g_t := (B_3 x_t)' s_t + \epsilon_t$, but they do not know the decomposition. Three unobservable and independent factors contribute to the signal g_t; Cognitive skills, motor skills and the iid productivity shock ϵ_t.

$$g_t = B_{3c} x_{ct} s_{ct} + B_{3m} x_{mt} s_{mt} + \epsilon_t, \tag{6}$$

where a worker's prior belief before signal g_t on each term is

$$\begin{aligned} B_{3c} x_{ct} s_{ct} &\sim N(B_{3c} x_{ct} \hat{s}_{c,t-1}, (B_{3c} x_{ct})^2 \sigma^2_{c,t-1}), \\ B_{3m} x_{mt} s_{mt} &\sim N(B_{3m} x_{mt} \hat{s}_{m,t-1}, (B_{3m} x_{mt})^2 \sigma^2_{m,t-1}), \\ \epsilon_t &\sim N(0, \sigma^2_\epsilon) \end{aligned} \tag{7}$$

A worker's prior beliefs on s_t are the expected skills given all information available up to period $t-1$, $\hat{s}_{t-1} = E(s_{t-1} | g_{t-1}, x_{t-1})$ and the variance for each skill is $\sigma^2_{c,t-1}, \sigma^2_{m,t-1}$. Workers update their beliefs on the cognitive and motor skills by Bayes rule.

By construction, the signal g_t is weighted by the task complexity x_t. Those who exert one skill more intensively than another gain more information on the skill that is used more intensively. For example, in the extreme case of $x_t = [1, 0]$, the occupation requires cognitive skill only, hence the worker will gain information on her cognitive skill but not motor skill by the signal g_t. For notational simplicity, let $\tau_{ct}, \tau_{mt}, \tau_\epsilon$ be $(B_{3c} x_{ct})^2 \sigma^2_{c,t-1}, (B_{3m} x_{mt})^2 \sigma^2_{m,t-1}$, and σ^2_ϵ respectively.

$$\begin{aligned} E(B_{3c} x_{ct} s_{ct} | g_t) &= B_{3c} x_{ct} \hat{s}_{c,t-1} + \frac{\tau_{ct}(g_t - B_{3c} x_{ct} \hat{s}_{c,t-1} - B_{3m} x_{mt} \hat{s}_{m,t-1})}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \\ &= \frac{\tau_{ct}(g_t - B_{3m} x_{mt} \hat{s}_{m,t-1}) + (\tau_{mt} + \tau_\epsilon) B_{3c} x_{ct} \hat{s}_{c,t-1}}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \end{aligned} \tag{8}$$

Hence, the posterior belief on cognitive skill given the noisy signal g_t is

$$(9) \qquad E(s_{ct}|g_t) = \frac{(\tau_{mt} + \tau_\epsilon)\hat{s}_{c,t-1} + \tau_{ct}\frac{(g_t - B_{3m}x_{mt}\hat{s}_{m,t-1})}{B_{3c}x_{ct}}}{\tau_{ct} + \tau_{mt} + \tau_\epsilon}$$

Similarly, the updated belief on motor skill is given by:

$$(10) \qquad E(s_{mt}|g_t) = \frac{(\tau_{ct} + \tau_\epsilon)\hat{s}_{m,t-1} + \tau_{mt}\frac{(g_t - B_{3c}x_{ct}\hat{s}_{c,t-1})}{B_{3m}x_{mt}}}{\tau_{ct} + \tau_{mt} + \tau_\epsilon}$$

The updated variance of cognitive skill is

$$(11) \qquad \begin{aligned} var(B_{3c}x_{ct}s_{ct}) &= \tau_{ct} - \frac{\tau_{ct}^2}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \\ &= \frac{\tau_{ct}\tau_{mt} + \tau_{ct}\tau_\epsilon}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \end{aligned}$$

$$(12) \qquad \begin{aligned} var(s_{ct}) \equiv \sigma_{ct}^2 &= \frac{1}{(B_{3c}x_{ct})^2} \; \frac{\tau_{ct}\tau_{mt} + \tau_{ct}\tau_\epsilon}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \\ &= \frac{\sigma_{c,t-1}^2(\tau_{mt} + \tau_\epsilon)}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \end{aligned}$$

Similarly, the updated variance of motor skill is

$$(13) \qquad \begin{aligned} var(s_{mt}) \equiv \sigma_{mt}^2 &= \frac{1}{(B_{3m}x_{mt})^2} \; \frac{\tau_{ct}\tau_{mt} + \tau_{mt}\tau_\epsilon}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \\ &= \frac{\sigma_{m,t-1}^2(\tau_{ct} + \tau_\epsilon)}{\tau_{ct} + \tau_{mt} + \tau_\epsilon} \end{aligned}$$

4.4 Bellman Equations

Combining all the components listed in this section, Bellman equation for a decision maker is formulated as:

$$\begin{aligned} V_t(m_t, \hat{s}_t, \sigma_t^2) &= \max_{c_t, x_{t+1}} u(C_t) + E_t(V_{t+1}(z_{t+1}, \hat{s}_{t+1}, \sigma_{t+1}^2)) \\ &= \max_{c_t, x_{t+1}} \frac{{C_t}^{1-\rho}}{1-\rho} + E_t(V_{t+1}(z_{t+1}, \hat{s}_{t+1}, \sigma_{t+1}^2)) \end{aligned} \tag{14}$$

$$\begin{aligned} s.t. \quad a_t &= z_t - c_t \geq 0 \\ z_{t+1} &= a_t + w_{t+1} \\ w_{t+1} &= B_0 + B_1' x_{t+1} - x_{t+1}' B_2 x_{t+1} + g_{t+1} \\ \hat{s}_{t+1} &= E_t(s_t | g_t) + A_1 x_{t+1} + x_{t+1}' A_2 x_{t+1} \\ \sigma_{j,t+1}^2 &= \frac{\sigma_{jt}^2 (\tau_{-j,t+1} + \tau_\epsilon)}{\tau_{j,t+1} + \tau_{-j,t+1} + \tau_\epsilon} + \sigma_{\eta j}^2, \;\; j = \{c, m\} \\ x_t &\in (0,1)^2 \end{aligned} \tag{15}$$

where $g_{t+1} \sim N(x_{t+1}' B_3 \hat{s}_t, (x_{t+1}' B_3 \sigma_t)^2 + (\sigma_\epsilon)^2)$, a_t is the end of the period t asset, amount of assets left after wage realization and the consumption decision in period t. And z_t is wealth, or cash-on-hand, available to use for consumption in the beginning of the period t. Individual workers choose current period consumption, and next period occupation to maximize expected lifetime utility.

Using backward induction, I numerically solve for the three choice variables which simultaneously satisfy the three first order conditions with respect to the each choice variable. In the final period T, the only choice that workers have is consumption. Workers exhausts their total wealth in the last period.

$$\begin{aligned} &V_T(m_T, \hat{s}_T, \sigma_T^2) = \max_{C_T} u(C_T) \\ &s.t. \; C_T = a_T + w_T \end{aligned} \tag{16}$$

One period before, the optimal choices for workers satisfy the following three first order conditions for C_{T-1}, $x_{c,T}$, $x_{m,T}$ respectively.

$$
\begin{aligned}
u'(C_{T-1}) =& E_t(V_T'(z_T, \hat{s}_T, \sigma_T^2)) \\
=& E_t(u'(C_T)) \quad \text{By Envelope Theorem}
\end{aligned} \tag{17}
$$

$$
E_{T-1}(\frac{\partial V_T}{\partial z_T}(B_{1c} - 2B_{2c}x_{cT} + B_{3c}s_{cT} + B_{3c}x_{cT}\frac{\partial s_{cT}}{\partial x_{cT}})) = 0 \tag{18}
$$

$$
E_{T-1}(\frac{\partial V_T}{\partial z_T}(B_{1m} - 2B_{2m}x_{mT} + B_{3m}s_{mT} + B_{3m}x_{mT}\frac{\partial s_{mT}}{\partial x_{mT}})) = 0 \tag{19}
$$

Equations (18) and (19) do not contain $\frac{\partial V_T}{\partial \hat{s}_T}$ and $\frac{\partial V_T}{\partial \sigma_T^2}$ terms because the beliefs of each skill do not have any effects in the final period T, since the only choice variable is consumption at T.

Finally, for the periods $t = 1, ..., T-2$, The first order condition with respect to $x_{c,t+1}$ is:

$$
E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}\frac{\partial z_{t+1}}{\partial x_{c,t+1}} + \frac{\partial V_{t+1}}{\partial \hat{s}_{t+1}}\frac{\partial \hat{s}_{t+1}}{\partial x_{c,t+1}} + \frac{\partial V_{t+1}}{\partial \sigma_{t+1}^2}\frac{\partial \sigma_{t+1}^2}{\partial x_{c,t+1}}) = 0 \tag{20}
$$

To show $\frac{\partial V_{t+1}}{\partial \hat{s}_{t+1}} = 0$, take a partial derivative of each side with respect to $\hat{s}_t$.

$$
\begin{aligned}
\frac{\partial V_t}{\partial \hat{s}_t} &= E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}\frac{\partial z_{t+1}}{\partial \hat{s}_t} + \frac{\partial V_{t+1}}{\partial \hat{s}_{t+1}}\frac{\partial \hat{s}_{t+1}}{\partial \hat{s}_t} + \frac{\partial V_{t+1}}{\partial \sigma_{t+1}^2}\frac{\partial \sigma_{t+1}^2}{\partial \hat{s}_t}) \\
&= E_t(\frac{\partial V_{t+1}}{\partial \hat{s}_{t+1}}\frac{\partial \hat{s}_{t+1}}{\partial \hat{s}_t}) \\
&= \frac{\tau_{ct}}{\tau_t}E_t(\frac{\partial V_{t+1}}{\partial \hat{s}_{t+1}})
\end{aligned} \tag{21}
$$

Since equation (21) holds for all periods,

$$
\begin{aligned}
\frac{\partial V_{T-1}}{\partial \hat{s}_{T-1}} &= \frac{\tau_{c,T-1}}{\tau_{T-1}} E_{T-1}(\frac{\partial V_T}{\partial \hat{s}_T}) \\
&= \frac{\tau_{c,T-1}}{\tau_{T-1}} E_{T-1}(\frac{\partial u(z_T - C_T + w_T)}{\partial \hat{s}_T}) \\
&= 0
\end{aligned} \tag{22}
$$

Similarly, $\frac{\partial V_{t+1}}{\partial \sigma^2_{t+1}} = 0$, and the first order condition for cognitive task choice $x_{c,t+1}$ is reduced to:

$$
E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}} \frac{\partial z_{t+1}}{\partial x_{c,t+1}}) = E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(B_{1c} - 2B_{2c}x_{c,t+1} + B_{3c}s_{c,t+1} + B_{3c}x_{c,t+1}\frac{\partial s_{c,t+1}}{\partial x_{c,t+1}})) = 0 \tag{23}
$$

Using the same process, the first order condition for motor task $x_{m,t+1}$ is again reduced to:

$$
E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}} \frac{\partial z_{t+1}}{\partial x_{m,t+1}}) = E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(B_{1m} - 2B_{2m}x_{m,t+1} + B_{3m}s_{m,t+1} + B_{3m}x_{m,t+1}\frac{\partial s_{m,t+1}}{\partial x_{m,t+1}})) = 0 \tag{24}
$$

Intuitively, beliefs affect the value function only through choices. Given a fixed occupational choice, having a higher or lower belief about ability does not change incomes, or the current and future utility values. Finally, the first order condition with respect to c_t is:

$$
u'(C_t) = E_t(u'(C_{t+1})) \tag{25}
$$

Assuming $A_{1c}, A_{1m}, A_{2c}, A_{2m} = 0$ for simplicity, equations (18), (19), (23), and (24) imply that the optimal choices of x_{ct+1} and x_{mt+1} for any $t < T$ are:

$$x^*_{ct+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(B_{1c} + B_{3c}s_{ct+1}))}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(2B_{2c}))} \tag{26}$$
$$x^*_{mt+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(B_{1m} + B_{3m}s_{mt+1}))}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}}(2B_{2m}))}$$

As long as $x^*_{ct+1}, x^*_{mt+1} \in (0,1)$, the optimal occupational task x_j is decreasing in B_{2j}, the cost of mismatch (equation 4), and increasing in B_{1j} and B_{3j}, which are the reward for the unit of tasks and the coefficient on the interaction term between the skill and the chosen task.

In the perfect information about the skills case, where the only remaining uncertainty in wage is in the idiosyncratic wage shock, the optimal occupational choices in equation (26) are simplified to:

$$x^P_{ct+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(B_{1c} + B_{3c}s_{ct+1})}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(2B_{2c})} = \frac{B_{1c} + B_{3c}s_{ct+1}}{2B_{2c}} \tag{27}$$
$$x^P_{mt+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(B_{1m} + B_{3m}s_{mt+1})}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(2B_{2m})} = \frac{B_{1m} + B_{3m}s_{mt+1}}{2B_{2m}}$$

In this case, the optimal occupation choices depend only on the workers true skill and the coefficients in the wage function.

If skills are unknown but workers are risk neutral, hence the utility function and value function are linear, the equation (26) is reduced to:

$$x^N_{ct+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(B_{1c} + B_{3c}E_t(s_{ct+1}))}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(2B_{2c})} = \frac{B_{1c} + B_{3c}\hat{s}_{ct}}{2B_{2c}} \tag{28}$$
$$x^N_{mt+1} = \frac{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(B_{1m} + B_{3m}E_t(s_{mt+1}))}{E_t(\frac{\partial V_{t+1}}{\partial z_{t+1}})(2B_{2m})} = \frac{B_{1c} + B_{3m}\hat{s}_{mt}}{2B_{2m}}$$

Again, in this case, the optimal occupations are determined only by the coefficients in the wage function and the mean of the skill beliefs.

However, if there are uncertainties and risk aversion, the first term inside the expectation in equation

(26), $\frac{\partial V_{t+1}}{\partial z_{t+1}}$, cannot be dropped, and the solution will depend on the curvature of the utility function with respect to consumption.

Workers face wage risks from two sources, from unknown skills s_c and s_m, and a transitory wage shock ϵ. Given any current belief $(\hat{s}_t, \sigma_t^2)$, the distribution of the unknown part of the future wage w_{t+1} is:

$$g_{t+1} \sim N(x'_{t+1} B_3 \hat{s}_t, (x'_{t+1} B_3 \sigma_t)^2 + (\sigma_\epsilon)^2) \tag{29}$$

Workers expect higher wages when $\hat{s}$ is higher, and larger variances when their belief is noisier (σ_t^2). Workers can control the size of the wage risk that comes from unknown skills through occupational choice. For example, in an extreme case, if a worker is particularly averse to wage fluctuations, she can minimize her wage risk (variance) to the minimum level σ_ϵ^2 by choosing $x_{c,t+1} = x_{m,t+1} = 0$. In doing so, however, this worker does not learn anything about her skills, and her expected skills in period $t+1$ will remain the same as in the current period.

As long as B_3 and $\hat{s}_t$ are positive, an increase in x_{t+1} will raise both the mean and the variance of the unknown part of the wage g_{t+1}, and the marginal effect on the known part of the wage is $B_1 - 2B_2 x_{t+1}$. If absolute risk aversion is decreasing, then the risk premium declines with respect to wealth, and hence the optimal occupational choice x_{t+1} will be higher when workers are rich, given the same beliefs. Therefore, workers choose occupations x_{t+1} not only by their wage coefficients and beliefs, but also by taking their risk preferences and wealth levels into account.

There is not a closed-form solution in this case, and it will be numerically solved by the algorithm introduced in the following subsection.

4.5 Algorithm to Solve the Model

Calculating the optimal amount of risk to take jointly with savings decisions in a multi-dimensional space is a difficult problem, which cannot be solved without complicated numerical computations. To speed up the calculation, I transform the three dimensions of continuous controls, consumption and two tasks, into a sequence of two optimization problems. For any given levels of assets and belief, I first find the optimal task choices that satisfy the first-order conditions in equations (23) and (24) simultaneously. With the optimal occupation policy function in hand, I calculate the expected future income for the five-dimensional grid of the state space, and this allows me to solve for the simple consumption choice, given current assets and expected income, the latter of which is derived using the occupation policy function.

To find the optimal tasks that simultaneously satisfy the two first-order conditions, I use Broyden's method (Broyden (1965)), which is an extension of the secant method of root finding to higher dimensions. The key idea behind Broyden's method is to calculate the whole Jacobians only once and to update using the secant information at other iterations. For root finding problems with simple Jacobians such as linear optimizations, Newton's method is more suitable for the point of view of the time effectiveness, because it converges in fewer iterations than Broyden's method. However, Newton's method requires repeated evaluation of the system of Jacobians for each iteration, Broyden's method can reach the optimum faster in complicated, non-linear optimization problems of this kind. I solve for the numerical optimums x_{ct} and x_{mt} that satisfy the first order conditions given the fixed grids of the five state variables, $(z_t, \hat{s}_{ct}, \sigma^2_{ct}, \hat{s}_{mt}, \sigma^2_{mt})$ (equation (26)). I obtain the policy functions for the two tasks choices by spline interpolation.

For the second step, I use the Endogenous Grid Method proposed by Carroll (2006). In this step, the objective is to find the value C_t, which has the same marginal valuation for each of the ends of the period asset value a_t using the first-order condition. And obtain z_t simply as the sum of a_t and C_t. As opposed to the usual solution methods that define ex-ante grids for z_t and then perform root-finding routines to find corresponding optimal C_t, the Endogenous Grid Method does not require a root-finding process; hence, it speeds up the numerical computation greatly. I numerically calculate the marginal expected value (EV_t') given the expected income w_t at the optimal occupation for each set of states $(z_t, \hat{s}_{ct}, \sigma^2_{ct}, \hat{s}_{mt}, \sigma^2_{mt})$ to find the optimal consumption C_t.

Finally, I evaluate the outside the grid chosen for solution by spline interpolation.

5 Estimation Results

I jointly estimate all the structural parameters; risk aversion, wage, and skill accumulation coefficients, initial wealth and initial skills, using Simulated Method of Moments. I simulate the career and savings choices of 12,085 workers (5 replications of 2,417 profiles observed in the NLSY79) using the observed individual backgrounds and work histories for 20 years after the labor market entry.

I construct 150 moments, including the mean of wages and task choices of each 1-year period after labor market entry for 20 years of the data. I use conditional moments for wages and task choices on the two different levels of educations: low if the final education is high school graduate or less, and high if some college or above. The full list of moments is provided in Appendix. The estimates $\hat{\theta}$ are defined by

$$\hat{\theta} = \arg\min_{\theta} \left\{ \sum_{k=1}^{K} \Big((M_k^d - M_k^s(\theta))^2 / \mathrm{Var}(M_k^d) \Big) \right\} \tag{30}$$

where (M_k^d represents kth data moment and $M_k^s(\theta)$ is kth simulated moment at the parameter value θ. I compute asymptotic standard errors following Gourieroux et al. (1993).

5.1 Identification

There are four sets of structural parameters to be identified: wage parameters, skill accumulation parameters, initial skill parameters, and the coefficient of relative risk aversion. Variation in individual histories in income and occupation data, together with observed differences in individual characteristics such as educational attainment, AFQT scores, and demographics, help identify the structural parameters.

Some complications for identification arise because workers choose occupations based on their unobservable beliefs about their skills, while their wages are then determined both by their task choices and their (unknown) skills. For identification, I assume rational expectations: i.e., individuals "know" the model, and they have model-consistent expectations (James (2011), Hincapié (2020)).

Wage parameters include the productivity signal, which consists of two different sources of uncertainty: skill uncertainties and the transitory wage shock (equation (4)). The model describes workers' learning process from the wage signals (equations (9), (10), (12), and (13)), and it specifies the structure of endogenous selections for occupations (equation (26)). The relative size of these two sources is key in determining the speed of learning (i.e., how quickly workers resolve the uncertainty about their own skill levels). The variance of the transitory shock distribution is identified by the histories of the workers' occupational choices. If skill uncertainties are substantial and the relative transitory shock is small, workers' responses (i.e., occupational mobility) to the wage signal must be very sensitive at the beginning of the life cycle and decrease quickly over time. If the variance of the transitory shock distribution is large, on the other hand, the signal is noisier, and it takes longer for workers to settle into their final occupations.

Skills evolve depending on a worker's current occupation and the idiosyncratic skill shocks that are assumed to be normally distributed (equation (5)). In contrast to the transitory wage shocks, the skill shocks have persistent effects on workers' productivity, since they change workers' skill levels permanently. Therefore, the correlations in residual income across time help to identify the size of the permanent skill shocks. In addition, the variance of workers' beliefs (their uncertainty about their own skills) converges to zero without the permanent skill shocks, implying that there will be no further occupational changes due to adjustments in beliefs after some periods of learning. Hence, the converged magnitudes of occupational mobility in the final periods identify the distribution of the permanent skill shocks.

The parameters for initial skills are identified by the correlation between a worker's initial occupational choice and his/her observable characteristics prior to labor market entry, including OLS parameters of the first occupational choices on the years of education and AFQT scores.

Finally, the risk-aversion coefficient ρ is identified by the occupational choices across the population with heterogeneous levels of wealth conditional on the annual wage signal. Workers choose future occupations, and occupations determine both the amount of risk in income and the accuracy of the future wage signal. With a positive wage signal, a worker will find that more difficult job will fit better for her unknown skills, however, the occupational change involves a risk, and the optimal amount of risk to take depends on the current wealth levels. As the difference between occupations' optimal policy functions for a risk-averse worker (equation (26)) and a risk-neutral worker (equation (28)) implies, optimal task decisions depend on the risk-tolerance of the agent; the occupational choices serve as evidence for the risk-aversion parameter.

5.2 Parameter Estimates

One of the key elements in this model is the stochastic process of wages, because that is a main source of uncertainty. Workers observe the productivity signal $g_t = B_{3c}x_{ct}s_{ct} + B_{3m}x_{mt}s_{mt} + \epsilon_t$ and update their beliefs to make a future occupational choice. The mean-zero idiosyncratic shock ϵ_t with standard deviation $\sigma_\epsilon = 0.5045$, reported in Row 5 in Table 4, guarantees that the signal is noisy. Therefore, workers cannot immediately pin down their skills after one year of work experience.

Table 4 shows that the reward for an additional unit of cognitive task, $B_{1c} = 14.5424$, is much higher than for motor task, $B_{1m} = 5.3875$.[2] Similarly, the coefficient on the interaction terms between task and workers' true skill is larger for cognitive tasks than for motor tasks, where $B_{3c} = 27.3512$ and $B_{3m} = 20.6370$. However, more interestingly, the cost for overshooting – that is, the cost of choosing a higher task complexity when one's true ability is low – is also much higher for cognitive tasks ($B_{2c} = 28.5607$) than for motor tasks ($B_{2m} = 19.2269$). Therefore, even though the compensation for cognitive tasks is higher than for motor tasks, having an occupation that requires high cognitive intensity may not be an attractive choice for risk-averse workers if the variance, which is uncertainty, in their cognitive skill belief is large.

Table 4 also reports the CRRA risk aversion coefficient, $\rho = 3.8666$, which is strictly greater than 0 and implies that workers are risk averse. Previous literature typically finds that the risk aversion coefficient for CRRA utility function is in the range of 1 to 5 (MaCurdy et al. (1990), Friedberg (2000)). The estimated risk aversion coefficient implies that when workers have uncertainty about their own

[2]This finding is consistent with the reduced-form results OLS and panel data fixed effect wage regressions in the Appendix.

skills, they will under-invest; they will choose lower-intensity tasks in situations with uncertainty, compared with situations in which their skills are perfectly known. Moreover, this inefficiency (due to uncertainty) will decrease with workers' wealth level.

Table 5 shows the estimates for the initial skills and the skill accumulation parameters. Cognitive skills accumulate almost linearly along the choice of cognitive task, where $A_{1c} = 0.0479$ and $A_{2c} = 0.0023$. The coefficient on the quadratic term of cognitive task, A_{2c}, is close to zero and not statistically significant. However, motor skill accumulation is concave in the motor task choice, where $A_{1m} = 0.0663$ and $A_{2m} = -0.0742$. Therefore, motor skill accumulates more as the choice of the motor task is larger. However, the marginal benefit of motor skill accumulation for choosing higher motor tasks diminishes as the motor task itself increases.

Idiosyncratic skill accumulation shocks, reported in the 3rd row in Table 5, accounts for the permanent shock in workers' productivity. The estimation results show that the standard deviation of the distribution of cognitive skill accumulation shock is 1.1081, while the standard deviation for the distribution of the motor skill accumulation shock is smaller, 0.0580.

I use two pre-labor market entry variables to determine initial skills: AFQT score and years of education. Both variables have a positive coefficient on the initial cognitive skill, where $H_{1c} = 0.0027$ and $H_{2c} = 0.1028$. On the other hand, both variables have a negative coefficient on the initial motor skill, $H_{1m} = -0.0012$ and $H_{2m} = -0.0519$, while the coefficient of AFQT on the motor skill is not statistically significant.

Row 8 of Table 5 reports the standard deviation of initial belief distribution for each of the skill dimensions. The standard deviation of the initial belief distribution for cognitive skill is 0.3212, while for motor skill it is 0.3866. Therefore, on average, workers start with a higher degree of uncertainty about their motor skills than their cognitive skills at the time of labor market entry. However, because of the higher uncertainty in cognitive skill accumulation (compared to motor skill accumulation), the uncertainty about motor skills resolves faster than it does for cognitive skills over the life cycle.

5.3 Model Fit and Implications

The model's prediction fits the observed data well overall. Figure 2 shows the life-cycle profiles of occupations for each education group, where "high education" means that the highest level of educational attainment was some college education or higher, and "low education" means that the highest level attained was high school graduate or lower.

Cognitive task choices in Figure 2 (a) show an increasing pattern over time for both education groups, with the large gap between the two. Rather surprisingly, the average cognitive task choice for the low-education group continuously rises over the life cycle, while cognitive task choices for the high-education group increase rather sharply during the early periods of their careers and then stay constant after about 10 years post-labor market entry. As a result, the gap between the cognitive task choices of the two groups slightly decreases over the life cycle, both within the data and in the model simulation.

Figure 2 (b), then, shows the life-cycle profiles of the average motor task for each education group. The low-education group always chooses higher motor tasks than the high-education group does, on average. However, average motor task choice does not show generally increasing patterns for both groups; rather, the groups' motor task choices stay constant overall. To be even more precise, the low-education group's average motor task choice increases slightly in the early periods of their careers, and decreases afterwards; for the higher-education group, their motor task choices continue to diminish slightly over all the years. Still, the overall changes in motor task choices are very small compared to the changes previously observed in cognitive tasks.

The life-cycle profiles of the hourly wage rates are presented in Figure 3. On average, both education groups receive higher wages as their experience in the labor market increases, though it should be noted that both data and simulation results show that the wage gap between the two education groups widens over life cycle. For example, the high-education group, on average, received about $3 more per hour compared to the low-education group in the first year of labor market entry. However, by the 20th year, the difference in hourly wages between the two groups is around $10. This is a common finding in the literature: wage gaps widen over time.

The findings in Figure 2 and Figure 3 suggest that the widening wage gap is not due to highly-educated workers choosing more complex tasks over time, but rather to the dynamic effect of on-the-job skill accumulation and learning. Highly-educated workers tend to choose higher cognitive tasks in the early periods of the life cycle, and through the skill accumulation and learning channels, this choice returns even higher cognitive skills for highly-educated workers.

While workers' true skills or beliefs are not observed in the data, the simulation results in Figure 4 show the progressions of the workers' beliefs about their skills; the means and the standard deviations of their beliefs. Figure 4 (a) reveals a widening gap in the means of the cognitive skill beliefs held by the two education groups, which drives the increasing wage gaps between the groups in turn. In the 20th year, the gap between the means of cognitive skills are about 0.25 larger compared to the first year. The 0.25 difference in the cognitive skills accounts for about \$4.78 of wage differences when $x_c = 0.7$. By contrast, the means of motor skill belief in both groups stays constant over the life cycle, while the low-education group shows small increments and the high-education group shows the opposite. Panels (c) and (d) in Figure 4 depict the standard deviations of beliefs over time. For both cognitive and motor skills the uncertainties drop for the first 8 years after entering the labor market and stay roughly constant after.

Figure 5 presents the hourly wage profiles across the quartiles of the wage distribution to show that the simulation results can replicate the dispersion of wages in data. The model prediction fits the data well, while the main discrepancy lies in the bottom 25% of the high-education group's income distribution. This accounts for the fact that simulated average wage profiles for high education in Figure 3 are slightly lower than the actual data show. As many previous studies on income inequality document, figure 4 shows that the variance of the income distribution among the highly-educated workers is much larger than the variance in the income distribution for the low-educated group.

5.4 Benefits of Learning

In this section, I analyze the relative importance of learning and skill accumulation for lifetime earnings. To separate the two dynamic effects of career choice, I simulated the model without any of the learning effects described in Section 5.3. Workers' skills can increase over time due to the on-the-job skill accumulation, and their beliefs change accordingly taking skill accumulation effects into account. However, they do not adjust their beliefs based on the productivity signals.

Figure 6 shows the average wage profiles of the workers with and without learning effects. The solid line represents the baseline model with both learning and skill accumulation, and the dashed line shows the simulation results without learning effects. Workers in the two cases start with the same beliefs and true skills, therefore the starting wages are the same. The overall increasing trends are also shown in both scenarios. However, when workers update their beliefs based on the productivity signals to allocate themselves into the occupations that fit better to their skills, they can earn much more. The baseline model with the learning effects show over 20% increases in the hourly wages 10 years after entering the labor market.

5.5 Under-Investment in Career Choice and Distributional Effects

In this section, using the estimated parameters, I analyze the distributional effects of under-investment in occupational choices, which I define as a gap between the utility-maximizing and income-maximizing occupational choices. The discrepancy between the two choices occurs due to a combination of risk aversion and informational friction (i.e., skill uncertainty). Even if workers expect to earn the highest income in a certain occupation given their beliefs about their skills, that occupation still might not be the optimal choice for them if the choice involves too much risk. Every risk-averse worker with uncertainty will under-invest. However, the size of this inefficiency will be larger for workers with low wealth if their risk preferences display decreasing absolute risk aversion; if they find the same amount of monetary loss more hurtful when they are poor compared to times they are rich.

Figure 7 graphically describes the inefficiencies in career choice. The dashed line represents perfectly informed workers' cognitive task choices for different asset levels given a fixed level of true skill. Not surprisingly, these workers' occupational choices only depend on their true skills because wages are only determined by skills, occupations, and idiosyncratic shocks. Therefore, workers who know their skill levels choose their occupations regardless of their current assets. Workers who have uncertainty about their skills, however, choose less cognitive tasks at all asset levels, given the same beliefs and the same true skills. Furthermore, we can see that the discrepancy between the utility-maximizing and income-maximizing choices is larger when current asset level is low. Figure 7 depicts only one dimension of the two skills, however, we can expect the same patterns for the motor tasks as well.

This inefficiencies due to the informational friction causes income losses through two channels: current wage drops and the loss in continuation values. The current wage drop is a direct result of choosing less-complex (i.e., easier) occupations. Given any fixed level of true skills, choosing any occupation other than the wage-maximizing occupation returns lower expected wages in the current period.

The second channel, the losses in continuation values, includes two different dynamic effects: skill accumulation and learning. The skill accumulation parameters in Table 6 suggest that workers in a more demanding occupation today will accumulate additional skills in both skill dimensions through on-the-job skill accumulation. This effect is more drastic for cognitive tasks than motor tasks.

Finally, workers learn more about their true skills when they exert their skills more. The productivity signal $g_t = B_{3c}x_{ct}s_{ct} + B_{3m}x_{mt}s_{mt} + \epsilon_t$ in equation (5) is weighted by the current occupation choice. For example, if a worker in the hypothetical occupation requires no cognitive skills at all, $x_{ct} = 0$, then this worker knows that g_t only consists of the productivity generated by her motor skills and the idiosyncratic shock, because she knows that she chose $x_{ct} = 0$. Hence, she will not learn any new information about her cognitive skills, and her updated cognitive skill (given the signal g_t) will be exactly the same as her previous expectation, as in equation (8). In addition, equations (9) and (10) demonstrate that the updated variances in skill beliefs become smaller as the chosen occupations themselves are larger. Therefore, workers who choose more intensive occupations will have more precious information about their true ability.

Hence, all of this to say, workers who choose a more intensive occupation today ultimately have bigger chances of finding themselves in even higher positions in the future, through these two dynamic effects.

To measure the inefficiency in career choices by wealth, I simulate occupational choices for over 20 years of the life cycle for perfectly informed workers and compare the resulting wage profiles with the baseline model from the estimated parameters by the first-period wealth levels. The first column in Table 6 shows the wage profiles for the low-wealth group. The direct effect of the wage drops in the first period is rather sharp for this group. Without skill uncertainty, choosing wage-maximizing occupations, on average, returns them \$9.5202 per hour, while the optimal choices under uncertainty yield only \$7.1563.

After the first period of work experience and updating beliefs, the differences between the two wage profiles are much smaller in each wealth group, compared to the first period. However, the difference remains higher in the lower-wealth group. For the higher-wealth group, by contrast, the differences between the two wage profiles are smaller, both in the first period and throughout the life cycle. We can observe only a small loss, \$0.2690 per hour, in the last period of the last two columns.

6 Policy Implication: Long-Run Effect of Baby Bonds

The Baby Bonds policy[3] has been proposed to reduce the wealth gap in the U.S. The main idea of the policy proposal is that a \$1,000 savings account would be opened at birth for every child in the U.S. and that children in low-income households would get an additional deposit of up to \$2,000 in their account each year. At the age of 18, each person would receive the account, which would be worth, at most, \$46,215, including interest. This fund could be used for wealth-building purposes only, such as pursuing a higher education, buying a house, or starting a business.

Using the model, I simulate the long-term effects of the Baby Bonds policy at two different levels of funds: annual supplement of \$1,000 and annual supplement of \$2,000, which would be worth \$23,948 and \$46,215, respectively, 18 years after the account openings. The proposal is intended to provide different amounts of annual supports from \$0 to \$2000 based on the household income levels during the account recipients' childhoods. However, given the limitations in NLSY79 data, I cannot observe

[3] American Opportunity Accounts Act, S.3766–115th Congress (2017-2018). Retrieved from https://www.congress.gov/bill/115th-congress/senate-bill/3766/

parents' income profiles from the moment that each person was born. Therefore, I assume that the same amount of funds is given to everybody, and I add the amount to workers' original initial assets holdings before labor market entry.

This model does not take into account the effects of achieving higher education or buying houses to increase one's assets, but rather evaluates the other side of this policy's potential impact: namely, its potential to provide a safety net for young adults to experiment and discover the career that best matches their skill level. Although this model can potentially be extended to the choices related to wealth-building activities, the current model does not include educational attainment, housing purchases, or business startup as endogenous choice variables. However, considering that all these choices can have positive long-term effects on a person's lifetime earnings – either through increases in initial skills, through additional wealth, or both – the simulation results reported in Table 7 can serve as a lower bound of the policy effects in alleviating income inequality.

Table 7 reports the effect of the additional (i.e., Baby Bonds) funds on income inequality over the life cycle. Each row in Table 7 presents the average annual income ratio between the bottom and the top of the first-period wealth distribution; bottom 50% to top 50%, bottom 25% to top 25%, and bottom 10% to top 10%. The results imply that the Baby Bonds policy would likely have a large, long-term effect in reducing income inequality over the life cycle. $1000 of annual funds for 18 years, which sum up (with interest) to $23,948 would increase the income ratio between the bottom 50% and top 50% from 0.5131 to 0.5157, while they would increase the income ratio between the bottom 25% and top 25% of initial wealth groups by 1.5%. And $46,215 (the upper limit of the Baby Bonds proposal) would increase the bottom 50% to top 50% income ratio by 3% and increase the bottom 25% to top 25% income ratio by 9.2%. The biggest impact is on the lowest decile of the wealth distribution. Because the inefficiencies due to skill uncertainty and risk preference are largest for the people with the lowest wealth, equal amounts of additional funds from the Baby Bonds have the biggest effects for them. $23,948 of funds will increase the income ratio between the bottom 10% and top 10% of wealth groups, from 0.2466 to 0.2561, and $46,215 will increase it to 0.3022.

7 Reduced Form Empirical Evidence on Learning

One important question to address when investigating the impacts that learning about one's true skill level has on a worker is whether learning indeed exists in the labor market. Although imperfect information about ability and learning has been widely discussed in the literature, there has not, in fact, been much empirical evidence to demonstrate its importance. One finding in the literature that suggests learning is the fact that job mobility decreases with age and tenure (Neal (1999)). In a recent paper, Arcidiacono et al. (2016) address the fact that those who receive wages in excess of their worker characteristics are more likely to stay in the same occupation.

Taking advantage of the continuous task complexity space in Yamaguchi (2012), I provide richer reduced-form evidence of learning. Specifically, I regress the log hourly wage on individual characteristics such as race, education, AFQT scores before labor market entry, occupation-specific experience, the cognitive and motor task requirements of their current job, and an interaction term between the two skill requirements. The "surprise" part of the wage is the residual. Workers cannot observe how much of the surprise is from their cognitive skills or motor skills, however, the relative intensity between cognitive and motor tasks in the current job is informative for workers to infer the source of this new information; Workers who use cognitive (motor) skills more in the current job will learn more about their cognitive (motor) skills than their motor (cognitive) skills.

I use a rough measure of the relative intensity in this section; occupations are either cognitive- or motor-task intensive. If the cognitive task requirement is higher than the motor task requirement in the current occupation, that occupation is called cognitive-task intensive (or "cognitive-intensive"); otherwise, motor-task intensive (or "motor-intensive"). Workers who have cognitive-intensive occupations are expected to learn more about their level of cognitive skill, and workers with motor-intensive occupations will learn more about their level of motor skill. Those who have a cognitive-intensive occupation and receive a positive signal, therefore, are expected to seek occupational moves that require greater cognitive intensity; those who receive a negative signal, by contrast, are expected to move "down" to a less cognitive-intensive job. Similarly, people with motor-intensive occupations are expected to seek future jobs with more motor tasks, once they find (by virtue of a positive signal) that they are capable in that type of skill; however, they would not be expected to move in the same direction with respect to cognitive-intensive work, since cognitive and motor skills reflect different dimensions.

I estimated log hourly wage equations with and without individual fixed effects. The estimated wage equation without fixed effects is:

$$\ln w_{it} = \alpha + X_{1it}\beta_1 + X_{2it}\beta_2 + u_{it} \tag{31}$$

from which the signal, or surprise, is calculated as

$$\begin{aligned} \text{signal}_{it} &= \hat{u}_{it} \\ &= \ln w_{it} - \alpha + X_{1it}\hat{\beta}_1 + X_{2it}\hat{\beta}_2 \end{aligned} \tag{32}$$

And a panel data regression with fixed effects and the signal are:

$$\ln w_{it} = \alpha_i + X_{1it}\beta_1 + u_{it} \tag{33}$$

$$\begin{aligned} \text{signal}_{it} &= \hat{u}_{it} \\ &= \ln w_{it} - \alpha_i + X_{1it}\hat{\beta}_1 \end{aligned} \tag{34}$$

where X_1 includes cognitive task x_c, motor task x_m, the interaction term of the two $x_c x_m$, occupational tenure, occupational tenure squared, years of the experience in the labor market and its squared, and X_2 includes AFQT score, years of education, race dummy.

The first column of Table 8 shows OLS estimates and the second column shows fixed effect estimates. When the observed occupation is the same as in the previous period and wage data is missing, the wage is assumed to be the same as the previous period. Both regressions show that there is a sizable difference in returns for cognitive and for motor tasks. The coefficients on the interaction term between the two are negative, and occupational tenure has positive effects on the log wage. The OLS regression controls for the AFQT score, years of education, and race dummy variables, and age.

Tables 9 and 10 show the regression results for cognitive and motor tasks, x_c and $x_m \in (0,1)$ respectively, chosen in $t+1$ on the tasks in period t using the wage residual predicted in the OLS regression and the fixed-effects panel data regressions, respectively. Therefore, Tables 9 and 10 show the direction of occupational movement in each of the two tasks. Dummy variable $D_{c,t} = 1$ indicates cognitive-intensive occupations where $x_{c,t} > x_{m,t}$.

The coefficients on signal$_t$ $\times$ $D_{c,t}$ in Tables 9 and 10 suggest that a worker who has received unexpectedly high monetary rewards will move up to jobs that require more of the abilities that the worker is currently using compared to the workers in the motor-intensive jobs. On the other hand, those who have received disappointingly low wages adjust themselves into new occupations that require different kinds of skills. Workers in cognitive-intensive sectors tend to move up into more intensive cognitive tasks as their wage residuals increase, while negative shocks may make them choose higher motor tasks instead.

8 Conclusion

In this paper, I present a structural model to analyze the role of wealth in individual career choices and lifetime earnings. Wealth provides a buffer for unexpected wage shocks that result from task-skill mismatch. Therefore, if workers' risk preference displays decreasing absolute risk aversion, those who have greater initial wealth are willing to endure more risk in order to find a better match, and they are more likely to find an occupation that fits their skills. Hence, wealth inequality could expand as a result of individual career choices even if every worker makes an optimal choice given their asset levels and their belief of about their skills.

I used the model to quantify the inefficiencies resulting from the informational friction and risk preferences across initial wealth levels and to simulate the effects of recently proposed Baby Bonds policy on income inequality over the life-cycle.

A number of studies document that the social cost of occupational and sectoral mismatch is not negligible (Jovanovic and Moffitt (1990), Sullivan (2010), James (2011)). Self-selecting into a better occupational match, therefore, could be an important way to achieve more efficient allocation in the labor market. Also, important human capital investments, such as higher education or job training, often occur within a context of considering occupational decisions. Moreover, changing occupations often involves considerable variation in permanent income, which has a direct impact on household consumption and welfare. This structural model of occupational choice given imperfect information and risk aversion could, therefore, be useful in understanding occupational choices and other important economic behaviors within the same context.

Table 1: Summary Statistics

Variable	Mean	S.D.	N
Demographics and pre-labor market characteristics			
AFQT	49.0790	30.1438	2417
Years of Education	13.2375	2.5353	2417
Hispanic	0.1105	0.3135	2417
Black	0.0732	0.2506	2417
Age at the labor market entry	21.1386	2.9532	2417
Initial assets (log)	5.5706	2.6359	503
Wage and Occupations			
Hourly wage	17.6904	10.4820	31157
Cognitive task	0.5018	0.2645	32774
Motor task	0.5291	0.2487	32774

Notes: Summary statistics for 1 pre-labor market entry characteristics: AFQT score, years of educational attainment, race, age at labor market entry, and log of initial money assets, such as savings account, and 2) Labor market outcomes: hourly wage rate and task choices, using NLSY79 data from 1979-2000. Wages and assets are in 2005 real US dollars.

Table 2: Summary of Panel Data

	Hourly Wage			Cognitive Task			Motor Task		
t	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
1	11.8307	5.7708	2284	0.4124	0.2506	2412	0.5313	0.2268	2412
2	13.0147	6.3849	2293	0.4392	0.2579	2413	0.5334	0.2312	2413
3	13.6809	6.9410	2325	0.4532	0.2619	2410	0.5366	0.2378	2410
4	14.6976	7.7216	2218	0.4678	0.2630	2321	0.5342	0.2391	2321
5	15.4201	7.8133	2045	0.4741	0.2596	2138	0.5406	0.2450	2138
6	16.5760	8.7493	1915	0.5005	0.2633	1999	0.5424	0.2513	1999
7	17.5429	9.3701	1804	0.5069	0.2638	1885	0.5395	0.2495	1885
8	18.1036	9.6347	1732	0.5113	0.2646	1812	0.5356	0.2502	1812
9	18.8502	10.2204	1647	0.5260	0.2620	1732	0.5321	0.2561	1732
10	19.1571	10.3732	1592	0.5243	0.2668	1658	0.5214	0.2506	1658
11	19.8593	11.1129	1525	0.5339	0.2633	1597	0.5248	0.2586	1597
12	20.2770	11.2297	1452	0.5418	0.2650	1523	0.5126	0.2549	1523
13	20.5647	11.4465	1381	0.5380	0.2671	1455	0.5156	0.2560	1455
14	21.1374	12.0812	1297	0.5465	0.2627	1372	0.5124	0.2573	1372
15	22.0019	12.9901	1200	0.5417	0.2648	1264	0.5182	0.2575	1264
16	22.0607	13.2607	1069	0.5442	0.2585	1155	0.5231	0.2588	1155
17	22.0675	12.8598	928	0.5549	0.2523	1000	0.5294	0.2631	1000
18	22.1935	13.0519	781	0.5550	0.2544	834	0.5247	0.2626	834
19	23.3223	13.7584	641	0.5610	0.2543	694	0.5186	0.2594	694
20	23.0562	13.0674	460	0.5554	0.2525	488	0.5198	0.2582	488
21	23.0695	11.9703	323	0.5538	0.2566	353	0.5120	0.2588	353
22	23.4238	12.3660	245	0.5436	0.2525	259	0.5147	0.2596	259

Notes: The means and the standard deviations, along with the number of observations for hourly wage, cognitive task, and motor task for years after labor market entry (t) from the NLSY79 data (1979-2000). Wages are in 2005 real US dollars.

Table 3: Transition Probabilities (%)

	$x_{i,t+1}$ Quintile				
$x_{i,t}$ Quintile	1	2	3	4	5
Cognitive Task					
1	57.63	20.85	11.41	8.82	1.29
2	16.61	60.55	10.12	10.52	2.18
3	10.70	11.32	61.11	13.21	3.67
4	5.04	6.75	7.24	71.97	9.00
5	1.29	2.17	3.73	18.04	74.76
Motor Task					
1	66.28	19.54	8.19	3.97	2.02
2	9.78	63.26	14.03	6.74	6.20
3	3.26	10.87	62.86	12.96	10.06
4	2.82	7.71	20.92	58.88	9.67
5	1.39	6.93	14.47	9.15	68.06

Notes: Total number of observations: 30,311

Table 4: Wage and Risk Preference Parameters

	Cognitive Task	Motor Task
	Wage equation	
Intercept (B_0)	1.5772	
	(0.8876)	
Reward for task (B_{1c}, B_{1m})	14.5424	5.3875
	(4.8243)	(2.1358)
Cost for mismatch (B_{2c}, B_{2m})	28.5607	19.2269
	(1.8486)	(2.9714)
Interaction with skill (B_{3c}, B_{3m})	27.3512	20.6370
	(6.5409)	(3.1327)
Std. of wage shock (σ_ϵ)	0.5045	
	(0.2970)	
	Risk preference	
CRRA coefficient ρ	3.8666	
	(1.6216)	

Notes: Parameter estimates and standard errors (in brackets) for the wage and risk preference parameters. Wage parameters are the determinants for hourly wage rate, and risk preference parameter represents the estimates for coefficient of the CRRA utility function.

Table 5: Skill Accumulation and Initial Skills Parameters

	Cognitive Task	Motor Task
	Skill accumulation	
(A_{1c}, A_{1m})	0.0479	0.0663
	(0.0151)	(0.0286)
(A_{2c}, A_{2m})	0.0023	-0.0742
	(0.0132)	(0.0418)
Std. of skill accumulation shock (σ_η)	0.1081	0.0580
	(0.0299)	(0.0136)
	Initial skills	
Intercept (H_{0c}, H_{0m})	-1.2069	1.4885
	(0.4533)	(0.0851)
AFQT score (H_{1c}, H_{1m})	0.0027	-0.0012
	(0.0012)	(0.0011)
Years of education (H_{2c}, H_{2m})	0.1028	-0.0519
	(0.0256)	(0.0087)
Std. of initial skill shock ($\sigma_{\eta 0}$)	0.3616	0.2737
	(0.0868)	(0.0376)
	Initial belief	
Std. of initial belief distribution (σ_{s0})	0.3212	0.3866
	(0.0967)	(0.0954)

Notes: Parameter estimates and standard errors (in brackets) for skill accumulation and initial skill determination. Both deterministic and stochastic elements in skill accumulation and initial skills are separately estimated for each skill dimension, cognitive and motor.

Table 6: Life-Cycle Wage Profiles:
Skill Uncertainty vs. Perfect Information

	Low Wealth		High Wealth	
t	Uncertainty	Perfect Info.	Uncertainty	Perfect Info.
1	7.1563	9.5202	16.2367	17.5047
2	8.7742	9.8499	17.6144	18.2162
3	9.5224	10.2101	18.3133	18.8986
4	9.8739	10.5946	18.9814	19.6188
5	10.1794	10.9723	19.6510	20.3458
6	10.5170	11.3464	20.3811	21.1048
7	10.8772	11.7775	21.1224	21.9584
8	11.2332	12.1921	21.9257	22.7993
9	11.6360	12.6394	22.7999	23.6678
10	12.0277	13.0984	23.7412	24.5417
11	12.3827	13.5565	24.6485	25.4814
12	12.7948	14.0349	25.4134	26.3307
13	13.2324	14.5196	26.1842	27.1259
14	13.5978	15.0078	26.9423	28.0135
15	14.1359	15.5245	27.9774	28.9257
16	14.5943	16.0212	28.8376	29.7521
17	14.5462	16.5922	29.3503	30.3500
18	15.5865	17.1150	30.0801	30.9501
19	16.0646	17.6465	31.0962	31.7995
20	16.5295	18.2040	31.5660	31.8269

Notes: Hourly wage profiles for workers with and without skill uncertainty, by first-period wealth level. Low (high) wealth indicates that initial assets are lower (higher) than the median.

Table 7: Baby Bonds Simulation

		Annual Supplemental Payment	
Annual Income Ratio	Original Sample	\$1000	\$2000
Bottom 50%/ Top 50%	0.5131	0.5157	0.5319
Bottom 25%/ Top 25%	0.3495	0.3548	0.3819
Bottom 10%/ Top 10%	0.2466	0.2561	0.3002

Notes: Simulation results for the effect of the proposed Baby Bonds policy on income inequalities over the life-cycle. Income ratios between average annual incomes by wealth groups are defined by 10%, 25%, and 50% of the first period of wealth distribution.

Table 8: Log Wage Regressions

	(1) OLS Log Wage	(2) FE Log Wage
$x_{c,t}$	0.818	0.629
	(0.0306)	(0.0279)
$x_{m,t}$	0.573	0.431
	(0.0379)	(0.0365)
$x_{c,t} \times x_{mt}$	-0.751	-0.733
	(0.0580)	(0.0579)
tenure$_t$	0.0759	0.0723
	(0.00333)	(0.00269)
tenure$_t^2$	-0.00307	-0.00257
	(0.000297)	(0.000244)
AFQT	0.00330	
	(0.000132)	
education	0.0267	
	(0.00157)	
hispanic	-0.0511	
	(0.00978)	
black	0.0204	
	(0.0105)	
constant	1.547	2.229
	(0.0280)	(0.0199)
N	31563	31563
adj. R^2	0.231	0.145

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $x_{c,t}$ and $x_{m,t}$ indicate the cognitive and motor tasks respectively. Tenure indicates occupational tenure. Education is measured by years.

Table 9: Career Transitions (OLS residuals)

	(1)	(2)
	$x_{c,t+1}$	$x_{m,t+1}$
$x_{c,t}$	0.720	-0.108
	(0.00731)	(0.00742)
$x_{m,t}$	-0.0714	0.698
	(0.00663)	(0.00673)
signal$_t$	0.0188	0.0220
	(0.00286)	(0.00290)
tenure$_t$	0.00530	-0.00155
	(0.00137)	(0.00139)
tenure$_t^2$	-0.000291	0.000225
	(0.000128)	(0.000130)
signal$_t \times D_{c,t}$	0.0196	-0.0425
	(0.00443)	(0.00450)
$D_{c,t}$	-0.00142	0.0256
	(0.00508)	(0.00516)
constant	0.179	0.204
	(0.00426)	(0.00432)
N	29309	29309
adj. R^2	0.548	0.480

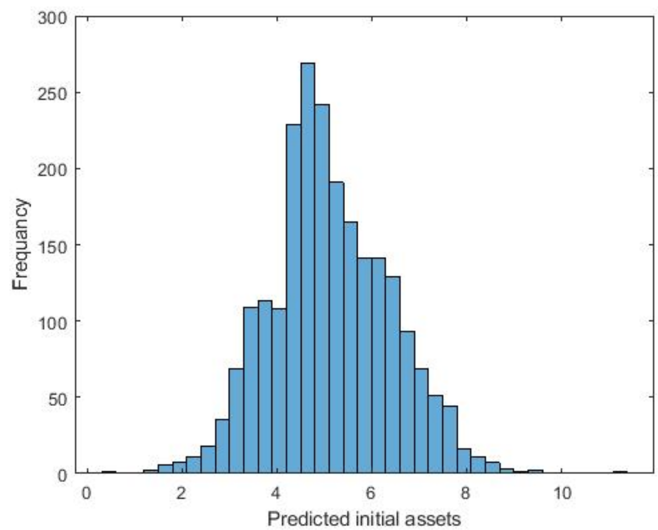
Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $x_{c,t}$ and $x_{m,t}$ indicate the cognitive and motor tasks respectively. Tenure indicates occupational tenure. Signal is the predicted residual from the OLS regression. Dummy variable $D_{c,t} = 1$ indicates $x_{c,t} > x_{m,t}$.

Table 10: Career Transitions (FE residuals)

	(1)	(2)
	$x_{c,t+1}$	$x_{m,t+1}$
$x_{c,t}$	0.716	-0.110
	(0.00730)	(0.00740)
$x_{m,t}$	-0.0675	0.698
	(0.00663)	(0.00672)
signal_t	0.0176	0.0147
	(0.00409)	(0.00414)
tenure_t	0.00528	-0.00163
	(0.00137)	(0.00139)
tenure_t^2	-0.000290	0.000226
	(0.000128)	(0.000130)
$\text{signal}_t \times D_{c,t}$	0.0234	-0.0556
	(0.00623)	(0.00631)
$D_{c,t}$	0.00211	0.0258
	(0.00508)	(0.00515)
constant	0.178	0.205
	(0.00426)	(0.00432)
N	29309	29309
adj. R^2	0.547	0.480

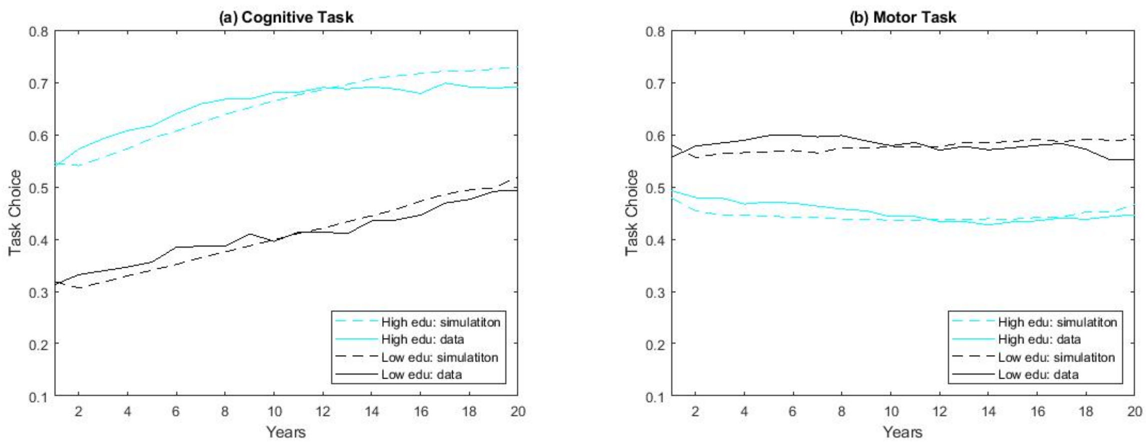
Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $x_{c,t}$ and $x_{m,t}$ indicate the cognitive and motor tasks respectively. Tenure indicates occupational tenure. Signal is the predicted residual from the fixed effect regression. Dummy variable $D_{c,t} = 1$ indicates $x_{c,t} > x_{m,t}$.

Figure 1: Predicted Log Initial Assets



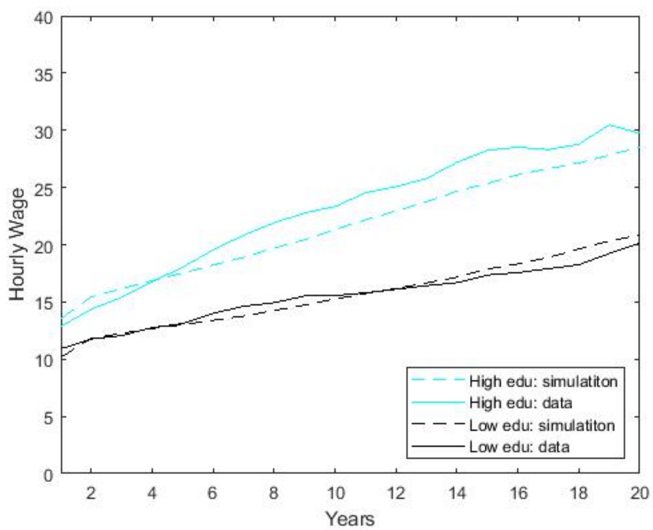
Notes: Predicted log initial assets at labor market entry. Predicted using the NLSY79 samples that have records of initial money asset holdings (such as a savings account), and years of education, AFQT score, first period wage rate and race.

Figure 2: Occupation Choice



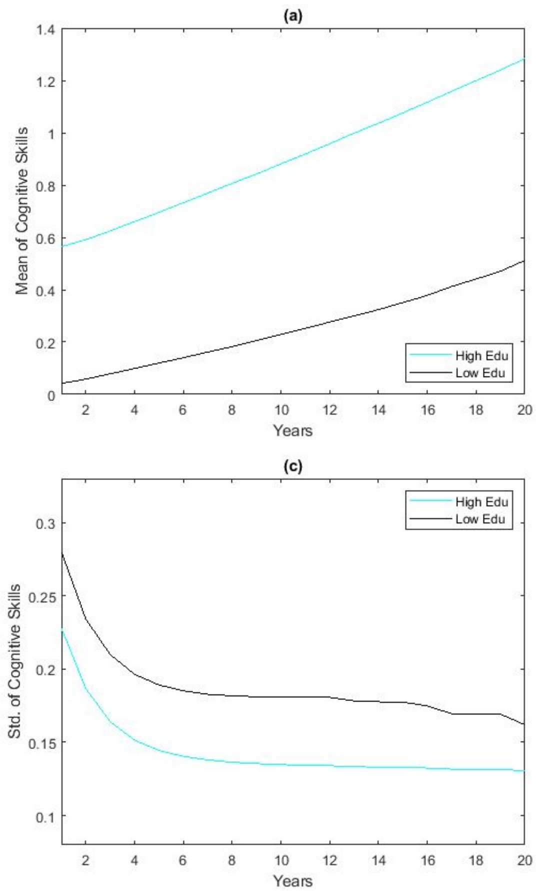
Notes: Average occupation choice profiles by education level. Education is indicated as high if the final education level attained is some college or above; low if the final level is high school graduate or lower.

Figure 3: Hourly Wage



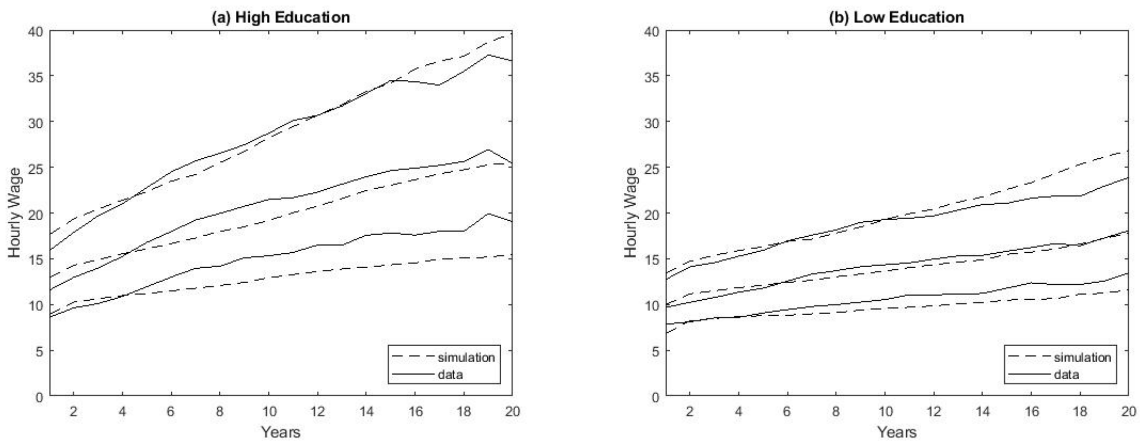
Notes: Average hourly wage rate profiles by education level. Education is indicated as high if the final education level attained is some college or above; low if the final level is high school graduate or lower.

Figure 4: Average Belief by Education Level



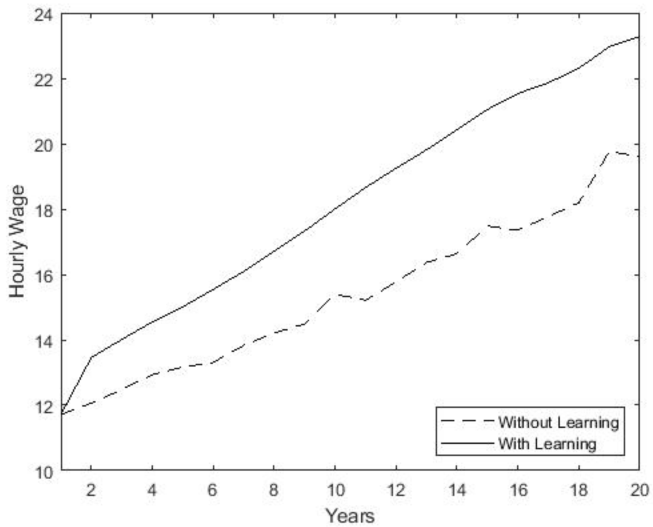
Notes: Average belief profiles of simulated data by education level. Education is indicated as high if the final education level attained is some college or above; low if the final level is high school graduate or lower. Panels (a) and (b) show the mean of noisy belief for cognitive and motor skill, respectively, and Panels (c) and (d) show the standard deviation of the belief regarding cognitive and motor skill over the life cycle.

Figure 5: Wage Distribution: Interquartile Range and Median



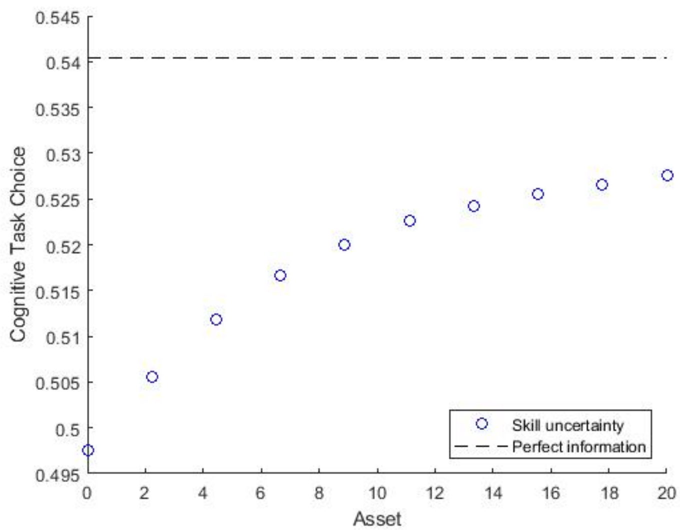
Notes: Hourly wage interquartiles and median, by education level. Education is indicated as high is the final education level attained is some college or above; low if the final level is high school graduate or lower.

Figure 6: Benefits of Learning



Notes: Dashed line indicates the average wage profiles for workers only with skill accumulation. Solid line represents the average wage profiles for the baseline model with both learning and skill accumulation effects. True skills and beliefs are fixed at the beginning of the life cycle.

Figure 7: Under-Investment in Occupational Choice



Notes: Dashed line indicates optimal cognitive task choices for a worker without skill uncertainty at each asset level; circles show optimal choices for an individual with skill uncertainty. True skills and beliefs are fixed at the population average, and assets are represented in the hourly dollar rate.

Appendix

1. Moments and Model Fit

Tables A1 - A6 display the full list of data moments, simulation moments, and normalized differences by the data standard deviation. The estimation was processed using 150 moments, including wage profiles, occupation profiles, OLS coefficients from the first period occupational choice regressions, interquartiles and the median of wage and occupations at the first, middle and the end of the panel. All moments except the OLS coefficients are education-specific.

2. Possible Extensions of the Model

2.1 Job Preferences

The recent literature has documented that non-pecuniary job preferences are one of the important factors for college major and occupational choices (Beffy et al. (2012), Wiswall and Zafar (2016)). Indeed, omitting job preferences from this model may have resulted in the biased estimators across different demographic groups. I did try a specification with random assignments of task preferences in the current model; however, the preference parameter estimates were not significant and close to zero. It is possible that such a result was due to the limitations of the data used in this paper; these data simply do not include good measures for individual-specific, pre-determined job preferences, so instead, I relied only on the individual work history (from the panel nature of the data) for the identification. Therefore, it is possible that the pre-determined differences in task preference may have appeared as differentials in workers' initial skills and beliefs. In other words, the model may have interpreted that people believe they are good at the tasks they like – which may be only partially true. Although the effect of job preferences is not the main focus of this article, the current model can be easily extended to include them. Such an extension would be especially interesting for studies concerning the comparison between broader demographic groups such as male and female, immigrants and non-immigrants, or domestic and foreign labor supply.

2.2 Moving Costs Across Occupations

Another interesting element worth discussing is the moving cost that occupational mobility entails. In particular, the moving cost is relatable if there are utility costs to moving; that is, if people prefer to stay in one occupation, or if there are monetary costs attendant to obtaining new skills for new jobs, such as retraining or education costs. In both cases, the role of wealth in occupational choice will be strengthened, and we are likely to see stronger distributional effects and income inequalities as a result.

2.3 Probability of Successful Match

This current model, furthermore, makes an important assumption about the worker-job match. Namely, it assumes that once workers choose occupations (though their earnings in the chosen occupations will depend on their true skills), they will certainly find a job within that occupation. In other words, the model assumes that there is no mismatch that results in unemployment. Once we introduce the probability of a match as a variable, along which the probability of the match declines as a worker's true skill level diverges further from (or sinks lower than) the skills required by a given occupational choice, then the income risks that workers face when they choose occupations will be larger, and again the role of wealth in occupational choice will be even more important. Such an extension can be interesting for studies that focus on unemployed job seekers; it may provide an interpretation for the role of wealth in unemployed people's occupational choices, and for the resulting match outcomes in terms of match qualities, new-job earnings, or unemployment duration.

Table A1: Average Cognitive Task Choices

	High education			Low education		
Moment	Data	Simulation	No. Diff.	Data	Simulation	No. Diff.
x_{c1}	0.5382	0.5453	0.0268	0.3128	0.3187	0.0313
x_{c2}	0.5725	0.5407	0.1220	0.3321	0.3068	0.1261
x_{c3}	0.5922	0.5565	0.1380	0.3393	0.3177	0.1043
x_{c4}	0.6078	0.5732	0.1379	0.3468	0.3298	0.0827
x_{c5}	0.6167	0.5932	0.0954	0.3560	0.3404	0.0758
x_{c6}	0.6399	0.6070	0.1329	0.3851	0.3517	0.1559
x_{c7}	0.6587	0.6231	0.1531	0.3860	0.3644	0.0981
x_{c8}	0.6675	0.6386	0.1289	0.3867	0.3756	0.0498
x_{c9}	0.6681	0.6516	0.0725	0.4101	0.3871	0.0997
x_{c10}	0.6803	0.6642	0.0726	0.3957	0.3987	0.0128
x_{c11}	0.6811	0.6760	0.0235	0.4132	0.4111	0.0091
x_{c12}	0.6902	0.6862	0.0188	0.4136	0.4209	0.0314
x_{c13}	0.6873	0.6958	0.0386	0.4105	0.4332	0.0961
x_{c14}	0.6912	0.7072	0.0745	0.4348	0.4446	0.0409
x_{c15}	0.6871	0.7121	0.1152	0.4365	0.4567	0.0818
x_{c16}	0.6788	0.7168	0.1778	0.4461	0.4724	0.1078
x_{c17}	0.6985	0.7216	0.1157	0.4688	0.4859	0.0707
x_{c18}	0.6915	0.7219	0.1465	0.4759	0.4942	0.0743
x_{c19}	0.6890	0.7251	0.1726	0.4916	0.4980	0.0260
x_{c20}	0.6911	0.7303	0.2068	0.4934	0.5195	0.1043

Notes: x_{ct} denotes the average cognitive task choice at year t. The education level is high is the final education is some college or above and low if the final education is high school graduate or lower.

Table A2: Average Motor Task Choices

Moment	High education Data	High education Simulation	High education No. Diff.	Low education Data	Low education Simulation	Low education No. Diff.
x_{m1}	0.4930	0.4788	0.0586	0.5563	0.5806	0.1166
x_{m2}	0.4795	0.4542	0.1012	0.5781	0.5563	0.1042
x_{m3}	0.4793	0.4465	0.1287	0.5835	0.5633	0.0962
x_{m4}	0.4672	0.4462	0.0831	0.5892	0.5655	0.1127
x_{m5}	0.4711	0.4437	0.1052	0.5983	0.5673	0.1455
x_{m6}	0.4688	0.4419	0.1029	0.5990	0.5696	0.1302
x_{m7}	0.4630	0.4412	0.0832	0.5958	0.5649	0.1403
x_{m8}	0.4577	0.4384	0.0724	0.5982	0.5748	0.1067
x_{m9}	0.4539	0.4376	0.0618	0.5878	0.5738	0.0616
x_{m10}	0.4429	0.4363	0.0250	0.5788	0.5771	0.0079
x_{m11}	0.4433	0.4358	0.0281	0.5849	0.5767	0.0355
x_{m12}	0.4337	0.4378	0.0158	0.5702	0.5769	0.0296
x_{m13}	0.4337	0.4379	0.0155	0.5774	0.5846	0.0319
x_{m14}	0.4272	0.4389	0.0441	0.5710	0.5849	0.0604
x_{m15}	0.4330	0.4382	0.0195	0.5748	0.5864	0.0510
x_{m16}	0.4350	0.4414	0.0247	0.5794	0.5910	0.0498
x_{m17}	0.4408	0.4421	0.0049	0.5830	0.5865	0.0146
x_{m18}	0.4379	0.4521	0.0538	0.5717	0.5909	0.0778
x_{m19}	0.4430	0.4513	0.0312	0.5512	0.5894	0.1556
x_{m20}	0.4469	0.4662	0.0732	0.5520	0.5902	0.1529

Notes: x_{mt} denotes the average motor task choice at year t. The education level is high is the final education is some college or above and low if the final education is high school graduate or lower.

Table A3: Average Wage Profiles

	High education			Low education		
Moment	Data	Simulation	No. Diff.	Data	Simulation	No. Diff.
w_1	12.8685	13.5471	0.1101	10.9078	10.1641	0.1530
w_2	14.3609	15.4483	0.1572	11.7585	11.8198	0.0114
w_3	15.4234	16.1520	0.0949	12.0679	12.2628	0.0359
w_4	16.7583	16.8638	0.0129	12.7630	12.6735	0.0142
w_5	18.0449	17.5391	0.0614	13.1049	13.0133	0.0148
w_6	19.5396	18.2174	0.1390	14.0068	13.4066	0.0876
w_7	20.8348	18.9004	0.1916	14.6254	13.7693	0.1175
w_8	21.9053	19.7030	0.1943	14.9364	14.2480	0.1005
w_9	22.7841	20.4276	0.2019	15.5288	14.7385	0.1041
w_{10}	23.3360	21.3279	0.1615	15.5691	15.2589	0.0447
w_{11}	24.5473	22.1730	0.1779	15.7922	15.7223	0.0101
w_{12}	25.0809	22.9545	0.1609	16.1115	16.1761	0.0087
w_{13}	25.7811	23.7409	0.1487	16.4218	16.6508	0.0293
w_{14}	27.2032	24.6760	0.1736	16.6811	17.1681	0.0614
w_{15}	28.2476	25.3580	0.1853	17.3535	17.8753	0.0613
w_{16}	28.5194	26.1170	0.1407	17.5642	18.3282	0.0975
w_{17}	28.3154	26.6518	0.1070	17.9081	18.8783	0.1144
w_{18}	28.7808	27.1366	0.1058	18.2509	19.6266	0.1356
w_{19}	30.4801	27.8233	0.1666	19.2351	20.3224	0.1017
w_{20}	29.7406	28.5372	0.0794	20.1576	20.8675	0.0643

Notes: w_t denotes the average hourly wage at year t. The education level is high is the final education is some college or above and low if the final education is high school graduate or lower.

Table A4: OLS Coefficients

Moment	Data	Simulation	No. Diff.
1st period cognitive task			
AFQT	0.0015	0.0013	0.0027
Years of ecuation	0.0408	0.0445	0.0210
Constant	-0.2007	-0.2312	0.0937
1st period motor task			
AFQT	0.0003	-0.0007	0.0099
Years of ecuation	-0.0163	-0.0174	0.0063
Constant	0.7295	0.7987	0.2097

Table A5: First Period Task Distribution

Moment	Data	Simulation	No. Diff.
1st period cognitive task			
25%	0.1962	0.18	0.0041
50%	0.4181	0.37	0.0362
75%	0.6235	0.607	0.0043
1st period motor task			
25%	0.3791	0.364	0.0044
50%	0.5099	0.525	0.0044
75%	0.6814	0.679	0.0001

Notes: The moments show the quartiles of the cognitive and motor task distributions.

Table A6: Wage Distribution

	High education			Low education		
Moment	Data	Simulation	No. Diff.	Data	Simulation	No. Diff.
1st period						
25%	8.9124	8.6013	0.0025	6.8034	7.8430	0.0457
50%	12.9184	11.5690	0.0479	9.9860	9.6683	0.0043
75%	17.6432	15.8756	0.0822	13.4223	12.7304	0.0203
11th period						
25%	13.2654	15.6897	0.0330	9.6952	11.0511	0.0381
50%	20.0306	21.6829	0.0153	14.0355	14.5443	0.0054
75%	29.4520	30.1093	0.0024	19.9243	19.4254	0.0052
20h period						
25%	15.4745	19.0259	0.0549	11.5918	13.4331	0.0279
50%	25.5153	25.4073	0.0001	17.7859	18.0764	0.0007
75%	39.6052	36.6334	0.0385	26.8394	23.8619	0.0728

Notes: The moments show the quartiles of the wage distribution at the 1st, 11th, and 20th year.

References

Antonovics, Kate and Limor Golan, "Experimentation and job choice," *Journal of Labor Economics*, 2012, *30* (2), 333–366.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College attrition and the dynamics of information revelation," Technical Report, National Bureau of Economic Research 2016.

Autor, David H and David Dorn, "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 2013, *103* (5), 1553–97.

_ , Frank Levy, and Richard J Murnane, "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 2003, *118* (4), 1279–1333.

Bacolod, Marigee P and Bernardo S Blum, "Two sides of the same coin us residual inequality and the gender gap," *Journal of Human resources*, 2010, *45* (1), 197–242.

Beffy, Magali, Denis Fougere, and Arnaud Maurel, "Choosing the field of study in postsecondary education: Do expected earnings matter?," *Review of Economics and Statistics*, 2012, *94* (1), 334–347.

Broyden, Charles G, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, 1965, *19* (92), 577–593.

Carroll, Christopher D, "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics letters*, 2006, *91* (3), 312–320.

Friedberg, Leora, "The labor supply effects of the social security earnings test," *Review of Economics and Statistics*, 2000, *82* (1), 48–63.

Gorry, Aspen, Devon Gorry, and Nicholas Trachter, "Learning and Life Cycle Patterns of Occupational Transitions," *International Economic Review*, 2019, *60* (2), 905–937.

Gourieroux, Christian, Alain Monfort, and Eric Renault, "Indirect inference," *Journal of applied econometrics*, 1993, *8* (S1), S85–S118.

Groes, Fane, Philipp Kircher, and Iourii Manovskii, "The U-shapes of occupational mobility," *The Review of Economic Studies*, 2015, *82* (2), 659–692.

Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer, "Multidimensional skill mismatch," *American Economic Journal: Macroeconomics*, 2020, *12* (1), 210–44.

Hawkins, William B and Jose Mustre del Rio, "Financial frictions and occupational mobility," *Federal Reserve Bank of Kansas City Working Paper*, 2016, (12-06).

Hincapié, Andrés, "Entrepreneurship over the life cycle: Where are the young entrepreneurs?," *International Economic Review*, 2020, *61* (2), 617–681.

Ingram, Beth F and George R Neumann, "The returns to skill," *Labour economics*, 2006, *13* (1), 35–59.

James, Jonathan, "Ability matching and occupational choice," *SSRN Scholarly Paper ID 1937879*, 2011.

Jovanovic, Boyan, "Job matching and the theory of turnover," *Journal of political economy*, 1979, *87* (5, Part 1), 972–990.

_ **and Robert Moffitt**, "An estimate of a sectoral model of labor mobility," *Journal of Political Economy*, 1990, *98* (4), 827–852.

Kambourov, Gueorgui and Iourii Manovskii, "Occupational specificity of human capital," *International Economic Review*, 2009, *50* (1), 63–115.

King, Allan G, "Occupational choice, risk aversion, and wealth," *ILR Review*, 1974, *27* (4), 586–596.

MaCurdy, Thomas, David Green, and Harry Paarsch, "Assessing empirical approaches for analyzing taxes and labor supply," *Journal of Human resources*, 1990, pp. 415–490.

Neal, Derek, "The complexity of job mobility among young men," *Journal of Labor Economics*, 1999, *17* (2), 237–261.

Papageorgiou, Theodore, "Learning your comparative advantages," *Review of Economic Studies*, 2014, *81* (3), 1263–1295.

Pavan, Ronni, "Career choice and wage growth," *Journal of Labor Economics*, 2011, *29* (3), 549–587.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford economic papers*, 1951, *3* (2), 135–146.

Saks, Raven E and Stephen H Shore, "Risk and career choice," *The BE Journal of Economic Analysis & Policy*, 2005, *5* (1).

Sanders, Carl, "Skill accumulation, skill uncertainty, and occupational choice," Technical Report, Working paper 2014.

Sullivan, Paul, "A dynamic analysis of educational attainment, occupational choices, and job search," *International Economic Review*, 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*, 1992, *107* (2), 439–479.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, human capital, and gender," Technical Report, National Bureau of Economic Research 2016.

Yamaguchi, Shintaro, "Tasks and heterogeneous human capital," *Journal of Labor Economics*, 2012, *30* (1), 1–53.