

# Skill Uncertainty, Skill Accumulation, and Occupational Choice\*

Carl Sanders

*Washington University in St. Louis  
and the Federal Reserve Bank of St. Louis*

October 9, 2012

## **Abstract**

Young workers make occupational transitions between similar occupations, e.g. graduate student to professor, as well as very different occupations, e.g. professional athlete to professor. While the first type are usually treated as normal moves and the second are downplayed, I argue that combining the theories of human capital accumulation and labor market experimentation can provide a way to explain both these types of changes. In addition, using data on the relative importance of each type of change, I can estimate the relative importance of the two theories for job mobility and wage growth. I construct a model where workers with a multi-dimensional skill bundle are able to invest in those skills and learn their skill levels by choosing occupations that require different amounts of tasks performed. Linking data on workers careers with occupational characteristics, I show how to separate human capital accumulation from experimentation and I estimate the model parameters. The results indicate that skill uncertainty explains approximately 30% of worker mobility across different task ratios but human capital accumulation drives wage growth over the life cycle.

---

\*Thanks to Christopher Taber, John Kennan, Salvador Navarro, Nicolas Roys, and Timothy Smeeding as well as many seminar participants for their help and suggestions.

# 1 Introduction

The occupational choices of young workers is usually discussed in the framework of workers choosing in order to develop their skills. Both anecdotally and in the data, the idea of workers moving from “unskilled” to “skilled” occupations is a standard way to think about the problem workers face. I will argue here that a more general view of the occupational choices of young workers can help us explain workers who make moves both towards higher skilled occupations but also moves to occupations that differ widely in tasks performed. In the data there workers who move over their career from “Counter Attendants, Cafeteria, Food Concession, and Coffee Shop Worker” to “Cooks, Restaurant” to “Food Service Managers,”<sup>1</sup> but also ones who move from “Engine and Other Machine Assemblers” to “Computer Operators” to “Maids and Housekeeping Cleaners” to “Inspectors, Testers, Sorters, Samplers, and Weighers.”<sup>2</sup> The first type of move is easily interpretable in the framework of someone who took jobs in order to develop food-specific skills and then became more productive and higher skilled across his career. However, the second move is tough to square with that same story.

In this paper I will show that a model with both a human capital accumulation motive and worker experimentation across occupations is necessary to explain both types of careers found in the data. To do this, I develop a model where workers have a multi-dimensional skill set, here cognitive and manual skills, and dynamically choose occupations that differ in the tasks performed. Workers do not have perfect information about their skills but need to spend time experimenting with tasks to gain additional information about their. Workers can choose different occupations to both invest in increasing their own skills and to learn about their unknown skill levels, and these two forces give the worker different incentives to take occupations. The model can explain why some workers who know their comparative advantage choose similar occupations over time to maximize their investment, while others who need to spend time to learn their skills find it best to take a series of different occupations.

In addition, using data on worker occupational choices linked with the tasks they perform in their occupations, I show that the model allows separation of the effects of skill un-

---

<sup>1</sup>Individual #39 in the NLSY79.

<sup>2</sup>Individual #2853 in the NLSY79.

certainty and experimentation from human capital accumulation. The relative importance of these effects is a long-standing problem in labor economics. In my setting, the multi-dimensionality of the tasks as well as seeing worker choices over time allows me to see not just that separations from occupations take place, but what choices workers make once they leave. This is significantly different than the usual firm-specific learning vs. human capital problem where the behavior of workers once they leave the firm is uninformative. Here, using the model intuition I can see how likely separations are to lead to workers choosing very different occupations, which suggests a poor match and learning, versus how likely they take similar jobs, which is indicative of a human capital-induced change.

More specifically, to test the predictions of the model and to estimate the model parameters, I use panel data from the National Longitudinal Survey of Youth 1979 on workers' wages, experience, and occupational choices over time. To determine the tasks they performed on the job, I link the occupational records with measures of tasks performed within occupations taken from the US Department of Labor O\*NET survey. In the data, there is significant worker mobility between occupations that have different task mixes. Also, workers switch between occupations that use different task mixes early in their career before switching between occupations which use similar ones. This is consistent with the predictions of the model if workers have uncertainty about their initial skill levels, and I use these data moments to help identify the estimated model.

In the end I am able to use the model to predict how worker mobility and wage growth would change in a world where there was no skill uncertainty or no human capital accumulation. The thought experiment I run in my counterfactuals is to inform workers about their true skills at the beginning of their careers and simulate how different their career paths would be. I find that if you measure occupational mobility in terms of changes in the ratio of tasks performed, there is 30% less occupational mobility. On the other hand, overall worker welfare does not decrease significantly (around 1%) because in the case of uncertainty experimentation is a way to hedge risk.

## **2 Previous Literature**

The literature of skill-specific human capital looks at more general forms of human capital than the “firm-specific” / “general” distinction that is typically made. In these models human

capital is through a vector that has different possibilities for substitution across different areas of the labor market. For example, the Roy (1951) model can be thought of as allowing for a 2-dimensional vector of skills where each skill is only used in one sector. Later extensions of the Roy model such as Heckman and Sedlacek (1985) allowed for a general vector of human capital that could be utilized at different rates across sectors.

With new data sets that allowed the observation of multiple types of individual ability, such as test scores for cognitive and non-cognitive ability, and data about the types of tasks workers perform in jobs, there is an emerging literature that studies the importance of different types of skills for labor market outcomes. Among these are Heckman et al. (2006), Poleshchuk and Robinson (2008), Abraham and Spletzer (2009), Borghans et al. (2010), and Robinson (2010), which look at the returns to different types of skills and the movement of workers across different tasks when they switch jobs. Models where human capital is occupation-specific can be seen as a case of the skill vector literature as well. The dynamic programming model of Keane and Wolpin (1997) allowed for investment into two different types of occupational human capital, white collar and blue collar, which could not be substituted with each other. Kambourov and Manovskii (2009) show that empirical evidence suggests that what has commonly assumed to be returns to firm-specific tenure are actually primarily explained by occupational-specific tenure, which is evidence that occupation-specific capital plays a large role in wage growth.

The majority of previous papers on multi-dimensional skills use static models. This paper contributes to a recent literature that extends the idea of multi-dimensional skills to a dynamic decision framework. Cunha et al. (2010) estimate a dynamic model of parental investment in children, where adult outcomes for the children depend on the levels of both cognitive and non-cognitive skills. Yamaguchi (2012) estimates a dynamic model of endogenous learning by doing with a multi-dimensional set of skills where workers choose which skills to invest in. My model is very similar to Yamaguchi's, but I allow for workers to ask based on beliefs about their skills and to rationally experiment across different tasks in order to learn them.

I also extend the literature of worker learning to a more general setting. The analysis of models where workers learn about their productivity began with Jovanovic (1979), who modeled the match between a worker and a firm as an experience good, so individual workers would need to spend time with a firm before determining whether it was a good match or

not. Much of the recent work on extending the Jovanovic model has emphasized other types of matches that the worker may be uncertain about. For example, one literature has modeled match effects between workers, firms, and also occupations. Miller (1984) and McCall (1990) study the optimal path of occupational mobility if workers can choose from risky or safe occupations but need to spend time in the risky occupations in order to learn their true productivity, and Papageorgiou (2012) and James (2011) estimate models of occupational uncertainty directly. More recently, Groes et al. (2009) use a model of uncertainty about the occupational match to explain the empirical U-shape of occupation mobility. Neal (1999) and Pavan (2011) use models with career matches, where a career consists of a series of occupations that are good substitutes for one another in terms of tasks performed and skills required.

My paper extends this literature modeling workers as being uncertain about their underlying skill levels rather than matches between firms, occupations, or careers. If one considers a specific firm, occupation, or career as consisting of a group of tasks that require worker skills, as workers learn about their skills it will look the same as learning their match for any given firm, occupation, or career. The difference is that skill uncertainty leads to a set of substitution possibilities that these other models do not consider: in a worker-firm match model, if the worker chooses to leave the firm the knowledge of the value of that particular match is worthless at other firms. The same holds across different occupations or careers. With workers learning about underlying skills, the fact that a particular match with one occupation went sour has information that is useful in differing amounts across new occupations. If the tasks are identical in the “new” occupation and it is just a different name, then the information that the worker’s relevant skill levels were low has useful information at the second occupation. On the other hand, if the new occupation is completely different, there may be no information from the current match.

The other paper to analyze skill uncertainty and experimentation is Antonovics and Golan (2012). They create a model where workers can endogenously take lower wages in order to learn about their skills, and they study the predictions of the model and show it can fit some stylized facts from panel data on occupational choice, occupational characteristics, and wages. Their model does not allow for skill accumulation, and their focus is primarily theoretical. My contribution to their paper is to provide a model with human capital accumulation as well as skill uncertainty and experimentation and to attempt to empirically

distinguish the two factors. In addition, I allow the worker's problem to be derived from a full equilibrium model and estimate the parameters of the model, allowing me to look at simulations of counterfactuals.

This paper is to my knowledge the only one to analyze the question of differentiating empirically between skill uncertainty and skill accumulation. However, this question is closely related to empirically distinguishing worker-firm match effects from increasing productivity due to firm-specific human capital. In both cases, looking at the relationship between wages and tenure (or in the case here, amount of time spent using skill) is not informative. This is because workers who have learned they are a bad match will have abandoned the firm, so the selection effect is that wages will be increasing over tenure. Here, looking at the cross-section of individuals who have spent a long time in a given occupation, their wages will be higher than those who have just entered the labor market, but it is unclear whether this is selection or skill accumulation. Farber (1999) and Nagypal (2007) have studied methods of decomposing wage growth over time within a firm to these influences. My solution is different from theirs because I use data not just on which occupations workers chose, but also how similar those occupations are in tasks.

### **3 Model**

The model augments the literature on skill-specific human capital accumulation with a learning motive. In contrast to models with only skill uncertainty or only skill accumulation, including both allows the model to show the differences the two motives imply for observed behavior. In particular, models with only skill uncertainty cannot match the fact that over time workers tend to move from lower task levels to higher ones. On the other hand, models with only skill accumulation cannot explain moves across different task mixes to jobs where accumulated skills would be less productive.

I believe this is the first model to attempt to separate the effects of endogenous human capital accumulation from endogenous information acquisition. The basic model is closely related to that of Heckman and Sedlacek (1985) extended to a stochastic framework with a dynamic selection equation for workers. Yamaguchi (2012) implemented a similar model empirically, although his model does not include skill uncertainty. Antonovics and Golan (2012) study a version of this model without skill accumulation in a single-dimensional frame-

work. But none of these models is capable of analyzing how to separate the effects of human capital growth from uncertainty and learning.

Time is discrete and runs from 1 to retirement at given age  $T$ . At labor market entry, a worker is endowed with two-dimensional skill bundle  $S^* \equiv (C^*, M^*)$  where  $C^* \geq 0$  and  $M^* \geq 0$  where the values are drawn from log-normal population distributions

$$\log C^* \sim N(\mu_C^*, \sigma_C^{2*}), \log M^* \sim N(\mu_M^*, \sigma_M^{2*}).$$

However, the worker does not know their true skills, but have beliefs about them denoted by  $C_0$  and  $M_0$ :

$$\log C_0 \sim N(\mu_{C_0}, \sigma_{C_0}^2), \log M_0 \sim N(\mu_{M_0}, \sigma_{M_0}^2)$$

where  $\mu_{C_0}$  and  $\mu_{M_0}$  are drawn from distributions such that  $E[\mu_{C_0}] = C^*$  and  $E[\mu_{M_0}] = M^*$ . Different workers will have different initial beliefs about their types, and they will be on average correct. Over time the individuals may gain better information about their skills leading to changes in the  $\mu$  and reductions in the  $\sigma^2$ .

In period  $t$  the state variables for the worker are the current beliefs about his skill vector, which is summarized by the random variables  $\log C_t \sim N(\mu_{C_t}, \sigma_{C_t}^2)$  and  $\log M_t \sim N(\mu_{M_t}, \sigma_{M_t}^2)$ . The fact that beliefs in every period will be log-normally distributed is a consequence of the nature of Bayesian updating that will be discussed later. So the state vector  $B_t$  (for “beliefs”) in period  $t$  can be summarized as a 4-dimensional vector  $B_t = (\mu_{C_t}, \sigma_{C_t}^2, \mu_{M_t}, \sigma_{M_t}^2) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$  where  $\mu_{C_t}$  and  $\mu_{M_t}$  are the worker’s current expectation of his log skills and  $\sigma_{C_t}^2$  and  $\sigma_{M_t}^2$  summarize the amount of uncertainty he has about those skills.

The only control variable for a worker in each period is the choice of occupation. Occupations are indexed by a 2-dimensional vector of tasks performed,  $\tau \equiv (\tau_C, \tau_M)$  where  $\tau_C \geq 0$  and  $\tau_M \geq 0$  are cognitive and manual tasks performed respectively. Occupations are otherwise identical, so for workers the problem of choosing an occupation is equivalent to choosing a  $\tau$  vector. There is a continuum of occupations and every  $(\tau_C, \tau_M)$  is a feasible choice for any worker in any period, so the control vector  $O_t$  (for “occupations”) can be summarized by  $O_t = (\tau_{Ct}, \tau_{Mt}) \in \mathbb{R}_+^2$ . Occupations and firms have the same information set as workers, so there is no possible lying or disagreement about worker skills.

The flow payoff in period  $t$  for the worker is his wages  $W_t$ , which are a function of his current beliefs about his skill levels and his chosen tasks. The parametric form should be

general enough to allow for a component of wages that doesn't depend on skills as well as some relationship between tasks assigned and the marginal returns to skills. A simple form that captures this is:

$$W_t(B_t, O_t) = a_0 - a_1\tau_{Ct}^2 - a_2\tau_{Mt}^2 + a_3\tau_{Ct} \exp(\mu_{Ct}) + a_4\tau_{Mt} \exp(\mu_{Mt})$$

where  $(a_0, a_1, a_2, a_3) \in \mathbb{R}_+^4$  is a fixed vector of parameters known to the agents but unknown to the econometrician. Higher task intensity occupations receive a higher fixed penalty for performing them but reward skills more, which in the end will lead to positive assortative matching of worker skills to tasks. For fixed  $\tau$ , wages are increasing in beliefs about average skills, and for a fixed set of beliefs there is a unique positive occupation choice  $O_t^*$  that would maximize wages in that period. Within a period the worker is risk neutral, but with the dynamic nature of the model his uncertainty will affect his choices as shown below.

The Bellman equation for the worker's full problem can be written as

$$V_t(B_t) = \max_{O_t} W_t(B_t, O_t) + \beta V_{t+1}(B_{t+1}) \quad \forall t = 1, 2, \dots, T-1$$

$$s.t. B_{t+1} = \mathcal{B}(B_t, O_t, Z_t) + \mathcal{F}(O_t)$$

where  $\beta$  is the fixed known discount parameter, and  $\mathcal{B}$  is a Bayesian update term that describes how beliefs will change due to learning, and  $\mathcal{F}$  is a deterministic growth trend for skills that depends on the chosen occupation. The Bayesian update term also depends on a stochastic component  $Z_t$  to be discussed below. There is also an end-of-career condition that the last period the worker just solves the static problem

$$V_T(B_T) = \max_{O_T} W_T(B_T, O_T).$$

In any period the worker would choose his occupation to maximize his period wages except that tomorrow's beliefs  $B_{t+1}$  are a function of the current occupation choice  $O_t$ . The changes to the worker's beliefs in the next period comes through two channels, skill accumulation and learning.



### 3.1 Skill Accumulation

One way the worker's beliefs will change is by endogenous skill accumulation. I allow workers to choose occupations in order to invest in their skills at different rates. Given the choice of occupation  $\tau_t$  the worker's true log cognitive skills will deterministically increase by  $\mathcal{F}_C(O_t) = \tau_C^{\alpha_C} R_C$  and log manual skills by  $\mathcal{F}_M(O_t) = \tau_M^{\alpha_M} R_M$ ,  $\alpha_C, \alpha_M \in [0, 1]$ .  $R_C$  and  $R_M$  are parameters known to and common across workers but are unknown by the econometrician. The effect this will have on a worker's beliefs is just that the mean will be shifted up by  $\tau R$  between periods and the variance will remain unchanged.

For the estimated model, it is necessary that  $R_C$  and  $R_M$  change as the worker ages so that his wages don't grow at the same rate his whole career. In the estimation the  $R$  will be parametrized by  $R_{Ct} = R_C \times t^{-\gamma_C}$  and  $R_{Mt} = R_M \times t^{-\gamma_M}$ ,  $\gamma_C, \gamma_M \geq 0$ .

### 3.2 Endogenous Learning

In addition to the skill accumulation component, the worker will learn about his true skills over time. Given the choice of occupation  $\tau_t$ , between period  $t$  and  $t + 1$  the worker will receive two independent signals  $Z_{Ct}$  and  $Z_{Mt}$  that have the forms

$$Z_{Ct} = \log C_t^* + \frac{\varepsilon_{Ct}}{\tau_C}, \varepsilon_{Ct} \sim N(0, \sigma_{C\varepsilon t}^2)$$

and

$$Z_{Mt} = \log M_t^* + \frac{\varepsilon_{Mt}}{\tau_M}, \varepsilon_{Mt} \sim N(0, \sigma_{M\varepsilon t}^2)$$

where  $C^*$  and  $M^*$  denote the true values of the worker's cognitive and manual skills and the  $\varepsilon$  are independent of each other and everything else. The worker cannot see the components of the signals, just the final values of  $Z_{Ct}$  and  $Z_{Mt}$ , but knows the functional form and distribution of the  $\varepsilon$ .

Since the worker's current beliefs about his true skills is log normal and the noise  $\varepsilon$  are normally distributed, from the worker's perspective the  $Z_t$  are both normally distributed as well with parameters

$$Z_{Ct} \sim N\left(\mu_{Ct}, \sigma_{Ct}^2 + \frac{\sigma_{C\varepsilon t}^2}{\tau_{Ct}^2}\right)$$

$$Z_{Mt} \sim N\left(\mu_{Mt}, \sigma_{Mt}^2 + \frac{\sigma_{M\epsilon t}^2}{\tau_{Mt}^2}\right).$$

Choosing higher-task occupations does not change the expected value of the signal, but it does reduce the noise associated with it. This is why the learning is endogenous: the worker can choose the amount of information conveyed to him by choosing different occupations.

After observing the signal, the worker updates his beliefs using Bayes' Rule. The posterior beliefs on log skills will be normally distributed since the priors over  $\log C_t$  and  $\log M_t$  are normal distributions and the data  $X$  are also normally distributed. It is well known that if the current beliefs about cognitive skills are  $\log C_t \sim N(\mu_{Ct}, \sigma_{Ct}^2)$ , the updating rule the worker will use upon receiving signal  $Z_{Ct}$  is

$$\mu_{C(t+1)} = \omega_t Z_{Ct} + (1 - \omega_t) \mu_{Ct}$$

$$\sigma_{C(t+1)}^2 = (1 - \omega_t) \sigma_{Ct}^2$$

where the weight is

$$\omega_{Ct} = \frac{\sigma_{Ct}^2}{\sigma_{Ct}^2 + \frac{\sigma_{C\epsilon t}^2}{\tau_c^2}}$$

and symmetrically for manual skills. The weight placed on the signal is a function of the current uncertainty about the skills, the “noise” of the signal  $\epsilon_{Ct}$ , and the chosen occupation. Higher task requirements will make the signals more informative and thus lead to the worker putting a higher weight on it for future beliefs, as well as reducing future uncertainty.

### 3.3 Full Problem

Putting together the two components of belief transitions, using the notation above I can write the full worker Bellman equation as

$$V_t(B_t) = \max_{O_t} W_t(B_t, O_t) + \beta V_{t+1}(B_{t+1}) \quad \forall t = 1, 2, \dots, T-1$$

$$s.t. B_{t+1} = \begin{pmatrix} \mu_{C(t+1)} \\ \sigma_{C(t+1)}^2 \\ \mu_{M(t+1)} \\ \sigma_{M(t+1)}^2 \end{pmatrix} = \begin{pmatrix} \omega_{Ct} Z_{Ct} + (1 - \omega_{Ct}) \mu_{Ct} + \tau_C^{\alpha_C} R_C \\ (1 - \omega_{Ct}) \sigma_{Ct}^2 \\ \omega_{Mt} Z_{Mt} + (1 - \omega_{Mt}) \mu_{Mt} + \tau_M^{\alpha_M} R_M \\ (1 - \omega_{Mt}) \sigma_{Mt}^2 \end{pmatrix},$$

$$\omega_{Ct} = \frac{\sigma_{Ct}^2}{\sigma_{Ct}^2 + \frac{\sigma_{C\epsilon t}^2}{\tau_C^2}}, \quad \omega_{Mt} = \frac{\sigma_{Mt}^2}{\sigma_{Mt}^2 + \frac{\sigma_{M\epsilon t}^2}{\tau_M^2}},$$

$$Z_{Ct} \sim N\left(\mu_{Ct}, \sigma_{Ct}^2 + \frac{\sigma_{C\epsilon t}^2}{\tau_{Ct}^2}\right),$$

$$Z_{Mt} \sim N\left(\mu_{Mt}, \sigma_{Mt}^2 + \frac{\sigma_{M\epsilon t}^2}{\tau_{Mt}^2}\right).$$

The general problem as written makes the intuition difficult, so in the next section I will go over the two-period version of this problem which is more tractable and illuminating.

## 4 The Two-Period Problem

### 4.1 Second Period

In the final period the problem facing the worker is (suppressing subscripts)

$$V(B) = \max_O W(B, O) = \max_{\tau_C, \tau_M} a_0 - a_1 \tau_C^2 - a_2 \tau_M^2 + a_3 \tau_C \exp(\mu_C) + a_4 \tau_M \exp(\mu_M).$$

This is just a quadratic problem that is separable in the  $\tau$ , so the first order conditions quickly give the solutions

$$\tau_C^* = \frac{a_3}{2a_1} \exp(\mu_C)$$

$$\tau_M^* = \frac{a_4}{2a_2} \exp(\mu_M)$$

and plugging back in gives the second period value function

$$V^*(B) = a_0 + \frac{a_3^2}{4a_1} \exp(\mu_C)^2 + \frac{a_4^2}{4a_2} \exp(\mu_M)^2.$$

A key feature of the second period value function is that it is convex in each  $\mu$  and even  $\exp(\mu)$ , which is the true units for skills. This is due to the nature of the positive assortative matching: a worker whose skills increased by a small amount would get a linear increase in wages without choosing a new occupation; if he is also allowed to choose a new occupation he can improve his wages quadratically since he can also move to a new occupation with a higher marginal return to that new unit of skill. This will have implications for attitudes toward uncertainty in the first period, since workers would prefer to take a fair gamble in skills over their second period value function.

## 4.2 First Period

The first period problem (suppressing subscripts again) is

$$V(\mu_C, \mu_M, \sigma_C^2, \sigma_M^2) = \max_{\tau_C, \tau_M} a_0 - a_1 \tau_C^2 - a_2 \tau_M^2 + a_3 \tau_C \exp(\mu_C) + a_4 \tau_M \exp(\mu_M) + \beta E \left[ V'(\mu'_C, \mu'_M, [\sigma_C^2]', [\sigma_M^2]') \right]$$

$$s.t. \mu'_C = \omega_C Z_C + (1 - \omega_C) \mu_{Ct} + \tau_C^{\alpha_C} R_C,$$

$$\mu'_M = \omega_M Z_M + (1 - \omega_M) \mu_{Mt} + \tau_M^{\alpha_M} R_M,$$

$$[\sigma_C^2]' = (1 - \omega_C) \sigma_{Ct}^2,$$

$$[\sigma_M^2]' = (1 - \omega_M) \sigma_{Mt}^2,$$

$$\omega_C = \frac{\sigma_C^2}{\sigma_C^2 + \frac{\sigma_{C\varepsilon}^2}{\tau_C^2}}, \quad \omega_M = \frac{\sigma_M^2}{\sigma_M^2 + \frac{\sigma_{M\varepsilon}^2}{\tau_M^2}},$$

Going through this one equation at a time: the first simply is the standard Bellman recursive representation of the problem. The second and third equations say that the mean beliefs tomorrow are a weighted average of the signal plus the old mean beliefs and shifted up by the amount of human capital accumulation. The fourth and fifth equations describe how uncertainty decreases over time with the amount of information received, and the last equations

summarize the weights. Since the only uncertainty in the transitions comes from uncertainty over  $\mu'$ , using the same argument as Antonovics and Golan (2012) we can just integrate out over  $\mu'$  using the fact that it is an affine transform of  $Z$  which is itself normally distributed. It is simple to show that

$$\mu' \sim N(\mu + \tau^\alpha R, \omega\sigma^2)$$

where this is true for each skill with the appropriate subscripts.

Now it is the case that the weight placed on the cognitive signal,  $\omega_C$ , is increasing in the task intensity  $\tau_C$ , and symmetrically for manual skills. Then for a worker who is thinking about what his mean belief about his skill will look like tomorrow, if he chooses a higher level of tasks his next-period mean beliefs will have higher variance. But given the convexity of the second period value function, this increased risk makes him have higher expected utility in the second period if he faces more risk over  $\mu'$  today. The intuition is straightforward from the standard discrete multi-armed bandit arguments. If the worker discovers he is terrible tomorrow, he can mitigate this discovery by taking a job that has low marginal returns to skills and a higher fixed payment. On the other hand, if he discovers he is a very high skilled worker, he can move to the part of the wage schedule with very high marginal returns and make a better-than-linear increase in wages in that period. The existence of jobs with just a flat wage is the equivalent of the safe arm of a multi-armed bandit, and the high-marginal-returns jobs are the risky arms.

This can be analyzed in closed form by solving for the continuation value:

$$\begin{aligned} E \left[ V'(\mu'_C, \mu'_M, [\sigma_C^2]', [\sigma_M^2]') \right] &= E \left[ a_0 + \frac{a_3^2}{4a_1} \exp(\mu'_C)^2 + \frac{a_4^2}{4a_2} \exp(\mu'_M)^2 \right] = \\ & a_0 + \frac{a_3^2}{4a_1} E[\exp(2\mu'_C)] + \frac{a_4^2}{4a_2} E[\exp(2\mu'_M)] \end{aligned}$$

and now use the properties of the raw moments of log-normal distributions to get

$$a_0 + \frac{a_3^2}{4a_1} \exp(2\mu_C + 2\tau_C^{\alpha_C} R_C + 2\omega_C \sigma_C^2) + \frac{a_4^2}{4a_2} \exp(2\mu_M + 2\tau_M^{\alpha_M} R_M + 2\omega_M \sigma_M^2).$$

The marginal effects of  $\tau$  on this continuation value summarize how much the dynamic nature of the problem affects current actions. First, even though within next period the

worker is risk-neutral, for the reasons mentioned above different levels of uncertainty in the first period lead to different incentives to choose that task. Second, it is easy to see that there is an incentive to choose take higher level tasks in the first period than the wage-maximizing tasks in that period because the continuation value is increasing in  $\tau$  with respect to both  $R_C$  (higher human capital accumulation rates means investment is more valuable) and  $\sigma_C$  (more uncertainty means information is more valuable.)

Even in this two-period case there is no closed-form solution for the occupation choice in period 1, although there is a (generically<sup>3</sup>) unique solution.

## 5 The Magnitude/Specialization Decomposition

The standard way that skill-based models view the task bundle in an occupation is Cartesian coordinates; with two skills, the task requirements for occupation  $j$  are  $\tau_j = (\tau_c, \tau_m)$ . Here I will use the polar representation of the occupation's task requirements in the empirical work. The reason for this is that the polar coordinates  $\tau_j = (\theta, \|\tau\|)$  where  $\theta = \arccos\left(\frac{\tau_c}{\|\tau\|}\right)$  and  $\|\tau\| = \sqrt{\tau_c^2 + \tau_m^2}$  have a natural interpretation in terms of specialization and total levels of task requirements. The second coordinate,  $\|\tau\|$  or the distance from the origin of the  $\tau_j$  vector, is the total amount of task complexity the occupation requires. The  $\theta$ , the angle from the x-axis, is the specialization of the occupation in the cognitive skills task. A  $\theta = 0$  represents a worker who uses only cognitive skills, and  $\theta = \frac{\pi}{2}$  is a worker who uses only manual skills, regardless of their total magnitude of tasks. See Figures 1 and 2 for a graphical representation of this transformation.

The reason this decomposition is useful is because I interpret similar levels of  $\theta$  as similar "career" type of occupations. Consider the case of the career path of college biology major to medical student to resident to surgeon. It seems plausible that the mix of manual dexterity to cognitive ability required is approximately unchanged across this path. But in terms of magnitudes, the college student certainly has the occupation with the lowest total task intensity, with the surgeon having the most. I will interpret moves across different levels of  $\theta$  as moves across different "career ladders", and moves with similar  $\theta$  and higher  $\|\tau\|$  as moves up one career ladder. I provide a list of occupations with their  $\theta$  and  $\|\tau\|$  scores in Table 4.

---

<sup>3</sup>Non-concavity of the  $\omega$  function could potentially lead to multiple solutions, but in this setting is it not important.

While it isn't possible to get results for the general model, I can use the  $\theta / \|\tau\|$  decomposition directly in a static version of the problem. Recall that the static problem facing the worker is

$$V(B) = \max_O W(B, O) = \max_{\tau_C, \tau_M} a_0 - a_1 \tau_C^2 - a_2 \tau_M^2 + a_3 \tau_C \exp(\mu_C) + a_4 \tau_M \exp(\mu_M).$$

Factoring the right hand side,

$$\max_{\tau_C, \tau_M} a_0 - a_1 \left[ \tau_C^2 + \frac{a_2}{a_1} \tau_M^2 \right] + a_3 \tau_C \exp(\mu_C) + a_4 \tau_M \exp(\mu_M).$$

Without loss of generality I can rescale  $\tau_m$  to  $\hat{\tau}_m = \tau_m \times \sqrt{\frac{a_1}{a_2}}$ , and the term within the maximum is

$$a_0 - a_1 [\tau_C^2 + \hat{\tau}_m^2] + a_3 \tau_C \exp(\mu_C) + a_4 \sqrt{\frac{a_2}{a_1}} \hat{\tau}_m \exp(\mu_M)$$

and now I define  $\|\tau\| = \sqrt{\tau_C^2 + \hat{\tau}_m^2}$  and  $\theta = \arccos\left(\frac{\tau_C}{\|\tau\|}\right)$ , and then by definition  $\tau_C = \|\tau\| \cos\theta$  and  $\hat{\tau}_m = \|\tau\| \sin\theta$ , so this equation is

$$p_0 - a_1 \|\tau\|^2 + a_3 \|\tau\| \cos\theta \cdot \exp(\mu_C) + a_4 \sqrt{\frac{a_2}{a_1}} \|\tau\| \sin\theta \cdot \exp(\mu_M)$$

and the value function can be rewritten

$$V(B) = \max_{(\theta, \|\tau\|)} a_0 - a_1 \|\tau\|^2 + \|\tau\| \mu_C \left( a_3 \cos\theta + a_4 \sqrt{\frac{a_2}{a_1}} \frac{\mu_M}{\mu_C} \sin\theta \right)$$

This maximization problem in two variables can be solved by a two-step process: first, choose the value of  $\theta$  that maximizes the term inside the parentheses, then given that value choose the optimal magnitude. Intuitively, the worker can choose the specialization based on the ratio of their expected skills regardless if their skill levels are high or low, and then choose the magnitude based on their overall skill levels. The value of  $\theta$  that maximizes this is  $\theta^* = \arctan\left(\frac{\mu_m}{\mu_c} \sqrt{\frac{a_1}{a_2}}\right)$ , which is independent of  $\|\tau\|$ . If this were a repeated static model with human capital accumulation, that is, the model with the discount factor  $\beta = 0$ , then if  $\mu_1$  and  $\mu_2$  rise proportionally from skill accumulation the optimal choice of  $\theta^*$  remains the same over time, even as skills grow, but the optimal value of  $\|\tau\|^*$  increases. On the other hand, with uncertainty but no skill accumulation, there would be shocks to the relative values of

$\frac{\mu_m}{\mu_c}$ . This would lead to different values of  $\theta$ , but on average the fact that  $\mu_c$  and  $\mu_m$  are not systemically increasing would mean the value within the parentheses above would remain about the same, which would mean the optimal choice of  $\|\tau\|$  remained about the same. This is another way of saying that skill accumulation type moves are ones with similar  $\theta$  and changing  $\|\tau\|$ , while learning moves are moves with changing  $\theta$  and stable  $\|\tau\|$ . In Sections 5 and 6 I use this intuition to look at the evidence regarding switching across  $\theta$ .

## 6 Data

The data I use comes from merging the National Longitudinal Survey of Youth 1979 (NLSY79) with the U.S. Department of Labor Occupational Characteristics Database, O\*NET. The NLSY79 is a representative sample of US households that was administered yearly from 1979-1994 by the Bureau of Labor Statistics, and once every two years since. The Weekly Labor Status variables can be used to create a weekly history of workers and the firms they work for, along with some characteristics of the job such as the hourly wage, hours worked, occupation, and industry. Additionally, the NLSY contains demographic variables such as age, sex, race, education, marital status, and scores on the Armed Forces Qualification Test (AFQT).

I restrict the NLSY79 sample to white male high school graduates from the core sample who have left school and have a firm attachment to the labor force. Specifically, I define a full time worker as one who works 30 or more total hours in a week at any employers. While there is no obvious definition for attachment to the labor force, I use criteria similar to Neal (1999) and Yamaguchi (2012) where workers who work full time at least 20 of the past 24 weeks are considered attached to the labor force, otherwise I do not use those observations. The time of labor market entry is then the first period where the worker is considered firmly attached to the labor force after leaving school. Only using high school graduates without college experience is another possible sample restriction, but for now I use those workers because of sample size considerations. Additionally, I drop workers who are ever in military service, since it makes sense they will have significantly different career paths than non-military individuals. From these weekly observations, I create yearly variables that contain the average wage paid during the previous year, the actual labor force experience of the worker, and the time since labor market entry as defined above. In addition, I merge in variables from the O\*NET database of occupation characteristics.



The O\*NET database contains occupational characteristics for a wide range of occupations across the US economy, at varying levels of detail. O\*NET is an update of the older Dictionary of Occupational Titles (DOT), which was criticized for being a non-representative sample of occupations, for not following standard survey design, and for poor data quality.<sup>4</sup> The current design of O\*NET has two stages, a worker survey stage and a job evaluation stage. In the survey stage, a nationwide random sample of workers complete a survey consisting of 8 different questionnaires, such as “Skills,” “Education and Training,” and “Generalized Work Activities.” An example question from the “Skills” questionnaire is shown in Figure 3.

In the evaluation stage, job evaluation experts are given the responses to read and fill out an “Abilities” questionnaire that is supposed to reflect the skill sets of the average individual in the occupation. All means and variances from both stages are released to the public, but not the individual-level responses.

I use the scores recorded in the O\*NET database to construct objects that I interpret as the  $\tau$ 's in the model. In particular, if I assume I am looking for two  $\tau$  for each job, the cognitive and manual task intensity, I need to determine how to reduce the information in many responses on the survey down to two variables. To do this, I follow Yamaguchi (2012) and first group the different questions into similar types of questions, where one group are questions that are related to cognitive tasks and the other group are related to manual tasks. I then use Principal Component Analysis to derive a single factor for both groups. This procedure creates a score for cognitive and task complexity as a weighted average of the scores on the individual questions, where if two questions tend to have the same responses in all occupations they receive less weight. The main function of this procedure is to attempt to rank jobs in terms of their cognitive skill and manual skill usage. I take the scores and rescale them between 0 to 1, corresponding to the occupation that uses the least amount of that task in the data and the job that uses that task the most, respectively. Summary statistics for the O\*NET database is shown in Table 1, listings of some selected occupations and their task scores are shown in Table 2, and histograms of the distributions of characteristics across populations are shown in Figures 4 and 5.

Looking at Figure 3, the question is made up of two parts: the importance of the skill, and the level of the skill. While this could possibly have interesting identifying information, in practice they scores are almost perfectly correlated and there isn't much to gain by using

---

<sup>4</sup>See Miller (1980) for a critical review of the DOT by the National Research Council.

both for each question. Instead, I just use the “level” score for each variable. Quantitative results for the regressions below are unchanged using the other score.

I combine the NLSY79 and O\*NET scores by merging the data sets using a crosswalk between the occupation codes in the NLSY79 and the 2000 Standard Occupational Classification codes used in O\*NET. To create this crosswalk, I took the titles from the occupation codes in the NLSY79 and searched the O\*NET website for the corresponding title. In the majority of cases there was an exact match between titles, or a closely corresponding title. In the majority of non-matched cases, the occupation was one occupation in the NLSY codes but is two in O\*NET or vice versa. The unmatched observations are counted as missing data. In future versions I will attempt other ways to get around the measurement problem of these missing occupations and the combined occupations, but over 70% of occupations were matched. The O\*NET occupational classifications are more detailed than the NLSY occupation codes, containing up to 10 sub-classifications of some occupations. Compared to some studies of task-specific human capital such as Autor and Handel (2009), I have data at a higher degree of aggregation, which could induce problems if there are wide amounts of skill dispersion within occupations that I do not observe.

The last step is to create yearly  $\tau_c$  and  $\tau_m$  scores for each individual. Since workers often work many different jobs in a single year, one possibility is to just take the task scores from the job worked at the most. Instead I create a week-weighted average of the scores, so if an individual works 26 weeks at a job where  $(\tau_c, \tau_m)_1 = (.8, .2)$  and 26 weeks at a job where  $(\tau_c, \tau_m)_2 = (.4, .6)$ , the total score given the year will be  $(\tau_c, \tau_m)_{total} = (.6, .4)$ . This assumes that working half a year at a manually-intensive job and then half a year at a cognitively-intensive job is the same as working for a whole year at a job that is an average of both. This does not follow from the model, but if the amount of information and skill accumulation that happens within a year is small it can serve as an approximation.

After data cleaning and restricting the sample, I work with 1,794 individuals from the NLSY from 1978-2008 with an average length in the sample of 25.7 years. For the O\*NET sample, I am left with data from 440 occupations to merge into the NLSY, and after using the crosswalk I am able to create yearly average task scores for 35,208 of the 39,462 (89%) yearly observations where a worker was firmly attached to the labor market. For some occupations I was unable to match task scores, so these occupations are treated as missing when I create the yearly averages. If these occupations are missing data for some systematic reason this

may be a problem, but most of the non-matches are where the recorded occupation are general categories like “Teacher, Other.” Summary statistics for the yearly version of the NLSY are shown in Table 3.

## 7 Evidence for Worker Learning

In the first empirical section, I will discuss some stylized evidence that workers experiment to learn about their skills. This can be interpreted as additional evidence beyond that of Antonovics and Golan (2012), who have a closely related model without skill accumulation. Skill accumulation in my model complicates matters, since both learning and skill accumulation generate endogenous occupational moves, and both can explain the decreasing transitions between occupations over time. In the model section I argued that moves across different specializations are consistent with a model where workers have to learn their skills. In this section, I analyze the empirical distributions and changes of the specialization and magnitude variables both across and within workers to see if these patterns are consistent with the model.

In order to differentiate between skill accumulation and skill uncertainty, I use the task data to create the specialization variable  $\theta$  as a proxy for the career path that an occupation is in. As workers accumulate skills, I assume that it only moves them up the task magnitude measure  $\|\tau\|$  and use the measures of  $\theta$  to look at experimentation across different career paths. The rest of this section is organized as follows: first, I discuss the distribution of the specialization and magnitude statistics in the data and look at the rates of change across specializations. Second, to look for evidence of experimentation I adapt an argument from Neal (1999) to my context and argue that optimal search behavior should lead to workers making “complex” changes, that is changes between occupations that perform very different tasks, before making “simple” changes, which are job switches to occupations that perform similar tasks.

Overall sample summary statistics for the values of the specialization and magnitude measures can be found in Table 3. This table shows that even though the sample average values of  $\tau_c$  and  $\tau_m$  (also in the table) are similar, they tend to be used in bundles where manual tasks make up a higher proportion of the overall task bundle. Looking at the aggregate levels does not give much information about individual-level dynamics, however, which

are the focus of this section.

Table 10 reports the means, medians, and standard deviations for the variables  $\Delta\theta$  and  $|\Delta\theta|$ , the signed value and absolute value of a worker's change in specialization since the last period respectively. The mean of  $\Delta\theta$  is nearly 0, which indicates that workers tend to remain using a similar mix of tasks between periods. This is true even conditioning on switching occupations. The standard deviation of  $\Delta\theta$  decreases slightly over time, which shows that there is a decrease in unusually large or small moves over time. The size of moves, regardless of direction, is also decreasing over time looking at the median of  $|\Delta\theta|$ .

The dispersion of workers across  $\theta$  is not simply given by an initial distribution of workers across jobs, but worker mobility within a career plays a significant role. The values in Table 10 help to summarize the amount of switching within individuals over time, as opposed to just a fixed distribution of workers across skills. Looking at the standard deviation of  $\Delta\theta$ , the value is around half of the total sample standard deviation of  $\theta$ . Movement across  $\theta$  within workers over time is a significant aspect of their careers, not just their initial level of  $\theta$ .

Just showing that workers move across different  $\theta$  is not direct evidence that workers move systemically across occupations since observed movement could reflect random shocks or measurement error. I present two related pieces of evidence to support the claim that this movement is consistent with experimentation. First, I create a variable every year for whether there was a "simple change" or a "complex change" from the previous year. A "simple change" is defined as one where the change of the value of  $\theta$  from the previous year is less than the sample median of all the year-to-year changes, and a "complex change" is when the change is greater than that. Neal (1999) first used these terms as definitions of different types of worker moves. He defines simple changes as ones that involve a change of firm but not of career, and complex changes to be changes of firm and career. While our models are quite different and there are no firm-level effects in my model, the intuition that experimentation involves making complex moves before simple moves largely carries over here.

In my case, if the change in  $\theta$  every period was drawn from the empirical distribution in the sample, I can calculate the probability that complex changes would never follow simple changes given any total number of observations. I then look at the number of workers in my sample where they actually follow that pattern. The results are shown in Table 5. They indicate that across different total numbers of possible changes, workers tend to follow the "complex change first" / "simple change second" pattern more often than random chance

would suggest.

There are a couple things to note about the results from Table 5. First, the results are similar even conditioning on firm switches like Neal does. Second, the magnitude of the gaps between the data and random movement are about the same as Neal's for a small number of possible moves, but not quite as large as his effects for the numbers of possible moves from 7 and up. Still, the values are uniformly higher than the predicted probabilities from random movement. This is especially true for individuals with a high number of moves. If movement was random almost no one in the data set would follow the pattern that is observed for approximately 5% of them.

There could be a few problems with the above analysis that do not come directly from the model, but may still be important. For one thing, perhaps the declining number of complex changes is just because of decreasing occupation to occupation or firm to firm mobility. Since my skills data is only at the occupational level, if there was some other reason that occupational mobility declined, it would look the number of complex changes were declining even if workers would still switch across skill sets if they were to switch occupations. Similarly, it could just reflect a decrease in firm-to-firm movements. With fixed occupation-specific or firm-specific matches, which I do not include in my model, if the worker finds a good match and stays there it adds a significant number of 0's (which simple changes) to the changes in  $\theta$ , which would raise the chances of the pattern "complex change" / "simple change" even though it has nothing to do with skills. To try to check whether this is what is happening, I repeat the analysis above for two different conditions. First, I condition on there being a change in primary occupation between the years. This raises the median size of the change in  $\theta$  (because of the elimination of all the 0's), so I redefine a complex change to be the median change conditional on the change being positive and recreate the table. The second check I do is condition on firm changes, calculate the new median size of the  $\theta$  move, and repeat the analysis again. While the tables are not included here, the results are basically the same as the unconditional version of the analysis, although the sample sizes in some of the cells get quite small. For each total number of changes, the pattern "complex changes first" / "simple changes later" is systematically overrepresented in the data compared to random movement. Since the results for the conditional analysis are so similar, I use the unconditional data for the regressions I run below.

The second set of stylized facts come from controlling for individual-level heterogeneity.

One way the results above could be spurious is if individual workers differed according to their probability of making complex changes, perhaps just because they had preferences for a higher or lower amount of occupational mobility than others. If there are some types of workers who are less likely to make simple changes and others who are less likely to make complex changes, the pattern would tend to look like the results above, since there would tend to be clustering of one type of changes within a career, and any pattern of just one type of change is consistent with the two-stage search argument. To try to control for this individual heterogeneity, I run a fixed-effects regression of a dummy for a complex change on the total number of simple changes that had been made up to that point in the worker's career, while also controlling for the length of time the worker has been in the labor force. Formally, the statistical model is a linear probability model OLS regression with individual fixed effects

$$\begin{aligned} \text{ComplexChange}_{it} = & \beta_0 + \beta_1 \sum_{k=0}^{t-1} \text{SimpleChange}_{ik} + \beta_2 \times \text{TimeSinceEntry}_{it} \\ & + \beta_3 \times \text{TimeSinceEntry}_{it}^2 + u_i + \varepsilon_{it} \end{aligned}$$

With only skill accumulation, changes in  $\theta$  should be correlated with time since entry but not necessarily the previous number of simple changes. With uncertainty, the expected sign of the coefficient  $\beta_1$  is negative, which is to say that as the number of times the worker does not move to an occupation that uses significantly different skills increases, he is less likely to make that move in the current period. The coefficients for this regression are displayed in Table 6, and the results are that for every previous simple change the worker made, he is 4% less likely to make a complex change this period. This is true even conditional on the total time since labor market entry, which is evidence that complex changes become less likely over time not just because all job switching rates decrease. Table 7 shows the same phenomenon when using the absolute value of the magnitude of the change in  $\theta$  from the previous period as the dependent variable, and the summed amount of the absolute values of past changes and time since entry as the controls. Formally,

$$|\Delta\theta|_{it} = \beta_0 + \beta_1 \sum_{k=0}^{t-1} |\Delta\theta|_{ik} + \beta_2 \times \text{TimeSinceEntry}_{it} + \beta_3 \times \text{TimeSinceEntry}_{it}^2 + u_i + \varepsilon_{it}$$

While these coefficients are more difficult to interpret, the negative estimate for the  $\beta_1$  coefficient indicates that the more the worker moved earlier in his career, the less likely he is to move in the current period, as long as the individual-specific heterogeneity in moving type is controlled for with the fixed effects. If the regression is run without the individual-specific effect, the estimate for  $\beta_1$  is positive and significant. Adding the fixed effects helps to control for the fact that individuals who are more likely to make larger moves in each period are more likely to have a large sum of past moves, which would make it look like past moves caused current moves.

In summary, looking at changes in the  $\theta$  measure over a worker's lifetime is consistent with a model where workers need to experiment not just across different occupations or firm but also across different tasks in order to learn their skills. In the next section, I will use the model explicitly to determine the importance of skill uncertainty for mobility across  $\theta$ .

## 8 Model Estimation

While the discussion above shows that there are aspects of the data that are consistent with a model of worker learning, it does not help to quantify the effects of learning compared to skill accumulation. If the quantitative implications of endogenous learning are small it may not be worthwhile to consider it further. To quantify the effects of skill uncertainty on task mobility, I take the model above and estimate it directly from the data. Using the estimates, I can simulate counterfactual worker careers, changing the amount of initial skill uncertainty or rate of skill accumulation.

The full model can be written

$$V_t(B_t) = \max_{O_t} W_t(B_t, O_t) + \beta V_{t+1}(B_{t+1}) \quad \forall t = 1, 2, \dots, T-1$$

$$s.t. B_{t+1} = \begin{pmatrix} \mu_{C(t+1)} \\ \sigma_{C(t+1)}^2 \\ \mu_{M(t+1)} \\ \sigma_{M(t+1)}^2 \end{pmatrix} = \begin{pmatrix} \omega_{Ct} Z_{Ct} + (1 - \omega_{Ct}) \mu_{Ct} + \tau_C^{\alpha_C} R_C - \delta_C \\ (1 - \omega_{Ct}) \sigma_{Ct}^2 \\ \omega_{Mt} Z_{Mt} + (1 - \omega_{Mt}) \mu_{Mt} + \tau_M^{\alpha_M} R_M - \delta_M \\ (1 - \omega_{Mt}) \sigma_{Mt}^2 \end{pmatrix},$$

$$\omega_{Ct} = \frac{\sigma_{Ct}^2}{\sigma_{Ct}^2 + \frac{\sigma_{C\epsilon t}^2}{\tau_C^2}}, \quad \omega_{Mt} = \frac{\sigma_{Mt}^2}{\sigma_{Mt}^2 + \frac{\sigma_{M\epsilon t}^2}{\tau_M^2}},$$

$$Z_{Ct} \sim N\left(\mu_{Ct}, \sigma_{Ct}^2 + \frac{\sigma_{C\epsilon t}^2}{\tau_{Ct}^2}\right),$$

$$Z_{Mt} \sim N\left(\mu_{Mt}, \sigma_{Mt}^2 + \frac{\sigma_{M\epsilon t}^2}{\tau_{Mt}^2}\right).$$

One difference with the estimated specification is that I have added skill depreciation terms  $\delta_C$  and  $\delta_M$ . I do not observe any of the worker's state variables, since I do not have any data on their beliefs about their skills levels. This is the opposite of a standard utility-maximization problem where the state (e.g. the worker's income) is observed but the level of utility is not. From observation of the age, choice, and wage, I try to back out the unobserved state. These types of problems are referred to as Hidden Markov Models in the inference literature. I cannot use typical solution algorithms from the HMM literature since my problem has endogenous transitions between states, and typical HMM algorithms assume a fixed exogenous Markov transition matrix.

The parameters of interest are the initial distribution of worker skills, the distribution of their initial beliefs about their skills, the variance of the noise  $\epsilon$  terms that dictates the speed of learning, and skill accumulation parameters  $R_1$  and  $R_2$ , and the depreciation rates  $\delta$ . I will discuss the identification of these parameters below in the "Indirect Inference Moments" section.

Since there is no closed-form solution to the problem, I use numerical methods to solve it. I use backwards induction on the value function with Chebyshev interpolation of the value function in each period and numerical integration of the continuation values.

Because of the computational complexity of the problem, even solving the individual's problem is difficult, since the value function has 4 continuous states and there is a double integral with no closed form solution inside the continuation value. Estimating the model is even more computationally demanding, since I do not observe the states, but only the controls and wages. Instead of directly solving the model, I will estimate the model using Indirect Inference, which has had some success in estimating learning models such as Guvenen and Smith (2008). The advantage is that this model is easier to simulate from than estimate: simulating the data for one individual is orders of magnitude easier than calculating the like-



likelihood for that individual. The essence of indirect inference is to simulate data from the model such that certain aspects of the simulated data compare well with the same aspects of the true data set. Included in this general method are Simulated Method of Moments and Simulated Maximum Likelihood estimators. I will estimate parameters by matching a series of moments from the data with the simulated data. Another important part of indirect inference is that the ways the model matches the data should be driven by the identification of the model. Here that leads me to choose moments related to wage growth, occupational switching rates, and the relationship between changes in  $\theta$  and time in the labor force, all of which I used in claiming evidence for the model in the first estimation section.

## 8.1 Indirect Inference Moments

### 8.1.1 Wages

The model has clear implications for wage growth. Workers become more productive through skill accumulation over time, while also gaining more information about their true skills levels. The set of indirect inference moments I use are mean wage rates in years 1, 5, 10, and 20 after labor market entry. The sample values from the NLSY are shown in Table 8. While my sample is restricted slightly differently these results are consistent with Topel and Ward (1992), who find that wages approximately double over the first 20 years of workers' careers. Considering the effects of model parameters of the simulated moments, the effects of  $R_c$  and  $R_m$ , the skill accumulation rates, are obvious: higher levels lead to more wage growth because workers become more productive. The effects of other parameters on this moment are more subtle. For example, changes any parameters that lead to workers being more uncertain about their skills at labor market entry (such as increasing the variance of the initial signal) lead to higher levels of wage growth, whereas changes to parameters that lead to less learning during the career (such as increasing the variance of the noise  $\sigma_\varepsilon$ ) lead to lower wage growth.

The second set of wage moments I will look at are the variances of wages in year 1, 5, 10, and 20. The variances of wages serves as a proxy for the distribution of the underlying skills in the economy,  $\sigma_c^0$  and  $\sigma_m^0$ . If jobs were randomly assigned, the variance of wages could be calculated directly from the variances of skills, since  $w_j = p_j + \tau_c \mu_c + \tau_m \mu_m$  and the distributions of  $\tau_c$  and  $\tau_m$  are observed. Under self-selection, this is no longer the case,

since different  $\tau$  pairs will have different  $s_c, s_m$  distributions that are generated by worker choices. Still, the intuition that high variance of skills leads to a high variance in wages in the economy should be true, and with endogenous skill accumulation and endogenous learning, high productivity types may find it worthwhile to invest and learn and low productivity types may not, leading to growth in the spread of wages as well as the overall levels. Wage variances are included in Table 8.

### 8.1.2 Task Intensity

The model also has predictions about the levels and changes for task intensity in the economy. Because of the higher occupation-specific prices of low  $\tau$  occupations, when workers begin their career with low skill levels the wage-maximizing occupations are the ones with lower task intensities. As they accumulate skills, they move up the occupation ladder to higher  $\tau$  occupations. This is only exactly true in a repeated static model, however. With skill accumulation and learning, higher  $\tau_c$  and  $\tau_m$  jobs are more attractive because of the opportunity to accumulate additional skills and learn about skill levels, both of which will have a positive value in the future. It is possible that these effects could completely offset each other, such that the increasing  $\tau_c$  and  $\tau_m$  that come from productivity gains could be offset by decreasing  $\tau_c$  and  $\tau_m$  motives once skill accumulation and learning become less important. The variance of the  $\tau$ 's also has information about the distribution of  $s_c$  and  $s_m$  across the economy. The mean levels of  $\tau_c$  and  $\tau_m$  along with their variances in years 1, 5, 10, and 20 after labor market entry are shown in Table 8.

I use two more moments that contain the task intensity measures, the means and variances of task specialization  $\theta$  and total magnitude  $\|\tau\|$ . As I discussed above, I interpret the  $\theta$  as a measure of the worker's specialization, where  $\theta = 0$  is a worker who only uses cognitive skills and  $\theta = \frac{\pi}{2}$  for a worker who only uses manual skills, regardless of the total overall level. The  $\|\tau\|$  then represent the overall tasks level irrespective of the specialization. Since both of these measures are constructed as non-linear transformations of  $\tau_c$  and  $\tau_m$ , the information in these moments is similar to the sample means and variances of  $\tau_c$  and  $\tau_m$  I used as other moments. However, they do provide additional information, since an economy with a high average level of both  $\tau_c$  and  $\tau_m$  could have any average  $\theta$  for some relationship between the individual  $\tau_c$  and  $\tau_m$ . Another similar possibility for a moment would be use the simple correlation between  $\tau_c$  and  $\tau_m$ . The means and variances of the  $\theta$  and  $\|\tau\|$  in years 1, 5, 10,

and 20 after labor market entry are listed in Table 8.

### 8.1.3 Relationship Between Tasks and Wages

The relationship between wages received and the choice of occupation is critical for the model. For these moments, I use the coefficient estimates from a regression of wages on time since entry and its square,  $\theta$ , and  $\|\tau\|$ . My predicted coefficients would have positive values on time since entry and  $\|\tau\|$  due to both learning and skill accumulation, a negative value on time since entry squared since the rate of skill accumulation and learning both slow down over time, and no guess about the sign of the coefficient on  $\theta$ . From the summary statistics it is expected that the coefficient on  $\theta$  will be negative, since jobs that are intense in manual skills pay less and higher levels of  $\theta$  mean a higher relative use of manual skills. Formally, the moments are the estimated values of  $\beta_1, \beta_2, \beta_3$ , and  $\beta_4$  in the following regression:

$$w_{it} = \beta_0 + \beta_1 \theta_{it} + \beta_2 \|\tau\|_{it} + \beta_3 \times TimeSinceEntry_{it} + \beta_4 \times TimeSinceEntry_{it}^2 + \varepsilon_{it}$$

I have run this regression in a fixed effects model, and the results are similar enough that I use the OLS estimates as my indirect inference moments. The coefficients are displayed in Table 9. The estimates have the expected sign, but the coefficients on  $\theta$  and  $\|\tau\|$  are actually quite small, since the standard deviations for those measures are approximately .2, so getting the effect of a one standard deviation change requires dividing the coefficient estimates by 5. The results are statistically significant, but this regression is not meant to capture a causal relationship. In the model there are many factors that jointly determine  $w$ ,  $\theta$ , and  $\|\tau\|$  and which are also dependent on time. This is only meant to capture the statistical relationship between wages and the  $\tau$  measures in the data and then to compare that to the one from the simulation.

### 8.1.4 Distribution of Changes in $\theta$

The last set of moments use changes within individuals across time to try to differentiate skill accumulation from learning. My argument in the model focused on changes in  $\theta$  to determine when a move is because of skill accumulation and when it is because of learning. Large moves in  $\theta$  imply that individuals chose to look for something different in terms of specialization, since moves generated by skill accumulation will tend to make the next occupation

have a similar mix of tasks. One of the moments I use to look at this is the mean and standard deviation of changes in  $\theta$  (denoted  $\Delta\theta$ ) in years 1, 5, 10, and 21 after labor market entry.<sup>5</sup> These means and standard deviations are displayed in Table 10. While the means don't have any consistent pattern over time and are relatively small, the standard deviation of changes in  $\theta$  has a slight decrease over time from around year 15 until the end of the career. The intuition from the model is that variance in changes in  $\theta$  is driven by uncertainty: with skill accumulation, workers should remain in approximately the same type of occupation they begin their careers in.

There is evidence for this in simulations of the model. I simulate 10,000 workers' careers where the only occupation changes are due to learning, and then those same 10,000 workers (all with the same initial conditions) where they know their skills exactly instead move only because of skill accumulation. I tuned the simulation so that there would be similar wage growth across the lifetime of these workers, and then compared the values of the variance of  $\Delta\theta$  between the two simulations. While the variance of  $\Delta\theta$  is still non-zero in the skill accumulation setting, it is less than half the size of the simulation with learning instead of skill accumulation, which the other moments (including others that are not listed in that table) are quite similar. The values of the moments from both simulations are shown in Table 11.

In order to match the individual-level mobility dynamics from the data, I use the fixed-effects regressions from the first part of the estimation section. The next set of moments are the estimated parameters from the fixed-effects regression

$$|\Delta\theta|_{it} = \beta_0 + \beta_1 \sum_{k=0}^{t-1} |\Delta\theta|_{ik} + \beta_2 \times \text{TimeSinceEntry}_{it} + \beta_3 \times \text{TimeSinceEntry}_{it}^2 + u_i + \varepsilon_{it}$$

In the data, the coefficient on  $\beta_1$  was negative, which meant that individuals who had more mobility across  $\theta$  were less likely to move in the current period. If the data generated by the model matched these effects, it would be because the information value of a different type of job had decreased due to earlier information. Estimates of these coefficients are reported in Table 7.

---

<sup>5</sup>I use year 21 for this moment since the variance in year 20 is unusual. The trend over time is that the variance is constant until around year 15 and then decreases the rest of the career. However, in years 19 and 20 the value spikes to what it was in year 10, which would give the false impression it is constant over the career. For the first 15 years after entry the average variance is .38. The values from year 15 onwards: .36, .35, .38, .39, .35, .35, .34, .33, .30, .32.

## 8.2 Estimation Procedure

To estimate the model, I choose model parameters to minimize the objective function

$$g(\hat{\gamma}) = \arg \min_{\hat{\gamma}} (h(\hat{\gamma}) - m)' W_n (h(\hat{\gamma}) - m)$$

where  $m$  is a  $k \times 1$  vector of the moments from the data,  $h(\gamma)$  is the function that maps any  $l \times 1$  vector of the parameters into a  $k \times 1$  vector that contains the same moments from the simulation, and  $W_n$  is a positive definite weighting matrix. For each possible value of  $\gamma$ , to get  $h(\gamma)$  I solve the individual's problem to generate the optimal policy for those parameter choices, and then I simulate 10,000 worker careers and calculate the indirect inference moments summarized above. I then search over values of  $\gamma$  until I have reached a local minimum, and repeat from different starting values to try to ensure I can find a global minimum to the objective function. For the weighting function, I follow [?](#) and use the diagonal matrix that contains the inverse of the standard errors of the moments.

To solve the individual problem for each set of parameter values, I use standard backwards induction of the value function in each period. Beginning in the last period, I use Chebyshev interpolation to generate coefficients for a polynomial that approximates the value function in that period, then in the previous period I do the same thing, using the interpolated continuation value. Getting the expected continuation value requires a numerical integration, which I do using Gauss-Hermite integration. Because of the structure of the problem, when it comes time to simulate the careers of the workers, I can just use the interpolated expected continuation value instead of having to do the integral of the value function again. Because of this, simulating worker careers is very fast relative to getting the Chebyshev coefficients the first time through. Calculating the Chebyshev coefficients takes around 2 seconds per period, which simulating 10,000 workers takes around the time to interpolate one period. The largest use of time for calculation comes from the numerical integration in the expected continuation value.

## 9 Results

In this section, I summarize the estimation results and the model fit. Tables 12, 13, and 14 show the values of the data moments and the corresponding moments the model generates at the estimated parameters.

The estimated parameters are shown in Table 15. While the parameter values themselves do not have clear interpretations outside of simulations of the model, a few things can be seen directly. I estimate that cognitive skills are more widely dispersed across the population than manual skills, which reflects the fact that the wage gap between high and low cognitive jobs is larger than the gap between high and low manual task jobs. The skill accumulation rate is estimated to be almost 0 for manual type jobs. This is because workers do not move to higher manual intensity tasks on average, so in the model they must not be accumulating additional manual skills. This result is consistent with the estimates of Yamaguchi (2012), whose estimates are from a model with only skill accumulation.

The other parameters reported in Table 15 are the estimates of the pricing function coefficients. When workers are deciding between occupations the differences in prices are all they consider, so I show the partial derivatives of the price function in the table along with the estimated parameters. The results indicate that higher cognitive and manual skill occupations are more desired by workers in the model and the lower prices for those occupations are used to reduce the amount of workers in that occupation.

My estimates of the amount of information that workers possess at labor market entry are quite low: using the estimates of the variances of the initial signal from Table 15 implies that workers' precisions about their beliefs are only 5% below the population precision. Looking at the model fit for aggregate moments in Table 12, the model fit for wage means and variances is quite good. The variance of wages grows slightly faster in the model than the data. Similarly, the model fit for cognitive tasks is overall quite close. The model overestimates the average manual task levels in the economy, and they do not decrease quite as quickly as they do in the data. The main area where the model has trouble fitting the aggregates is in the variances of task choices. The simulated standard deviations are around half as large as those observed. The reason for this is that the model needs to have small variances of skills in order to match the standard deviations of wages, but then at this distribution of skills workers will all take very similar jobs. One potential reason for this is that I do not allow for

measurement error in tasks. Autor and Handel (2009) document a large amount of task dispersion even within occupational categories, so potentially workers could be misreporting their task bundle which would make it appear that they are more dispersed than they really are. Table 13 has a similar interpretation, the model cannot fit the fact that the magnitude of skills has a small effect on wages in the data, while it should be important in the model.

Along with aggregate moments, I use individual-specific moments related to task mobility. The second part of Table 13 and also Table 14 show the estimates from those moments. Similar to the aggregate variances, the model predicts less mobility across different specializations than is observed in the data. In Table 14, I show that the model accurately reproduces the pattern of the effect of the size of past moves on current moves that I discussed in Section 5 above.

Overall, the model can do a good job reproducing the aggregate moments of the data as well as the general trends in the individual-level time series. The largest amount of room for improving the fit lies in the fact that workers' task choices are too similar to each other and they do not move enough across different tasks in the simulation. The main way I plan to fix this problem in future work is by estimating a more general CES production function and allowing for a more flexible specification of the pricing function. The current estimates using a Cobb-Douglas technology impose an elasticity of substitution of 1 between cognitive and manual skills. Allowing for more general substitution patterns between the skills could lead to workers choosing a wider variety of tasks even with similar skill levels.

## 9.1 Counterfactuals

I have emphasized the role that skill uncertainty can play in explaining the fact that workers move across jobs of different specializations. In this section I will show that in the estimated model, skill uncertainty and experimentation play a major role in this phenomenon. To do this, I use the estimated model parameters and change those that govern the amount of skill uncertainty so that workers have perfect information about their skills, keeping all other parameters the same. I then simulate the model under both sets of parameters and look at the changes induced in the movement of workers across jobs with different task mixes. If a large amount of the movement across specializations would be eliminated with perfect worker knowledge, that is evidence that experimentation can play an important role in using

models of skill accumulation to explain task mobility.

From the model, there are three primary ways that workers can move across different specializations. First, if the skill accumulation rates are different, e.g. if cognitive skills accumulate faster than manual skills the increase in the level of cognitive tasks workers use is faster than manual tasks, and workers would not stay at the same angle from the x-axis as before. Second, the equilibrium pricing function could potentially drive worker to be more specialized at higher task levels than they would be at lower task levels. Third, workers could experiment across different specializations in order to learn their skill levels early on. The first two reasons are ways the standard skill accumulation model can explain mobility across different specializations, and the third is a contribution to the model. From the parameter estimates it is clear that the skill accumulation rates are different across skills and that the pricing function is not symmetric over skills. To determine how much mobility skill uncertainty explains, it is necessary to simulate the model.

The results from the counterfactual simulations are shown in Table 16. The first column shows the values of three measures of the magnitude of moves in the simulated data at the true parameter values. The first measure is the mean of  $|\Delta\theta|$ , which is the average value of a worker's move between the last period and the current one in the sample. The second measure is the median of the size of moves, and the third is the standard deviation of the size of moves. The mean and median are both measures of the overall tendency for workers to move to jobs where they perform tasks with different specializations than they use in their current job, and the standard deviation summarizes the fact that the size of moves differs across workers.

The second column in Table 16 shows the results of a simulation where the variance of the worker's initial signal is set to 0 and holding the rest of the parameters fixed. The thought experiment this corresponds to would be each worker being told their true skill levels as soon as they enter the labor market, and then they solve their lifetime problem where the only dynamic motive is skill accumulation. The results indicate skill uncertainty decreases the mean size of moves by 41%, the median of the size of moves by 30%, and also decreases the standard deviation of the size of moves by 23%. Intuitively, since workers no longer need to use skills they are uncertain about to learn about them, there is less of a need to move across different specializations. While there is not one statistic that summarizes the amount the distribution of moves changes by, it seems reasonable to summarize these results by saying



that skill uncertainty can explain about 30% of movement across different task specializations.

## 10 Conclusion

Workers entering the labor market are uncertain about their skill set. In this paper, I argue that worker skill uncertainty can help explain worker mobility across jobs where they perform different types of tasks, a fact that standard models of skill accumulation typically ignore. To do this, I combine a model of worker skill uncertainty and skill accumulation with panel data on workers' tasks over their career. Intuition from the model suggests that workers will move across jobs where they use tasks in different mixes when they are uncertain about their skills. Mobility patterns across different specializations in the data are consistent with the model. I estimate the parameters of the model and find that skill uncertainty accounts for approximately 30% of mobility across different specializations.

The model extends previously used models of skill-specific human capital accumulation to allow for skill uncertainty and worker experimentation in an equilibrium framework. Workers choose what tasks to perform each period, but instead of basing their decisions on their true skill levels they act on beliefs about those skill levels and experiment in order to gain information. Using the model, I show that the larger the skill uncertainty and value of experimentation, the more workers will move across jobs that use significantly different tasks. In the model, taking jobs where workers perform different tasks than they did before has the highest value early in the career, when they have not worked in similar jobs before. I also allow worker skills to change over time through skill accumulation, which comes from a learning-by-doing mechanism. The incentives for investment are stronger in occupations where the worker believes he is already high skilled.

To look at whether the data is consistent with the implications of the model, I combine panel data on worker careers from the National Longitudinal Survey of Youth 1979 with data on the tasks performed in occupations from the US Department on Labor O\*NET database. From this data I decompose the worker's chosen bundle of tasks into measures of "specialization" and "magnitude," and document that moves across different specializations are common in the data.

I then estimate the parameters of the model and simulate counterfactuals to determine

the empirical importance of skill uncertainty in explaining moves across different specializations. Results from the simulations indicate that if workers knew their true skill levels at labor market entry, the size of changes across specializations would be reduced by approximately 30%. This suggests that worker experimentation plays an important role in worker mobility across tasks.

For future research, I will look at counterfactuals that allow the wage structure to shift in equilibrium. In this paper, the counterfactuals are only partial equilibrium and hold wage contracts fixed, but if it were possible to estimate the parameters of the labor demand side of the market I could use this framework to analyze the effects of shifts on the amount of worker skill uncertainty on the equilibrium distribution of wages.

## References

- Abraham, K. and Spletzer, J. (2009). New evidence on the returns to job skills. *American Economic Review: Papers and Proceedings*, 99(2):52–57.
- Antonovics, K. and Golan, L. (2012). Experimentation and job choice. *Journal of Labor Economics*, 30(2):333–366.
- Autor, D. and Handel, M. (2009). Putting tasks to the test: human capital, job tasks and wages. *NBER Working Paper*.
- Borghans, L., Ter Weel, B., and Weinberg, B. (2010). Interpersonal styles and labor market outcomes. *Journal of Human Resources*, Forthcoming.
- Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.
- Farber, H. (1999). Mobility and stability: The dynamics of job change in labor markets. *Handbook of labor economics*, 3:2439–2483.
- Groes, F., Kircher, P., and Manovskii, I. (2009). The U-Shapes of Occupational Mobility. *University of Pennsylvania working paper*.
- Guvenen, F. and Smith, A. (2008). Inferring labor income risk from economic choices: An indirect inference approach. *Working Paper*.

- Heckman, J. and Sedlacek, G. (1985). Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market. *The Journal of Political Economy*, 93(6):1077–1125.
- Heckman, J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.
- James, J. (2011). Ability matching and occupational choice.
- Jovanovic, B. (1979). Job matching and the theory of turnover. *The Journal of Political Economy*, 87(5).
- Kambourov, G. and Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, 50(1):63–115.
- Keane, M. and Wolpin, K. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.
- McCall, B. (1990). Occupational matching: A test of sorts. *Journal of Political Economy*, pages 45–69.
- Miller, A. (1980). *Work, jobs, and occupations: A critical review of the Dictionary of Occupational Titles*. Natl Academy Pr.
- Miller, R. (1984). Job matching and occupational choice. *The Journal of Political Economy*, 92(6):1086–1120.
- Nagypal, E. (2007). Learning-by-doing versus learning about match quality: Can we tell them apart? *Review of Economic Studies*, 74(2):537–566.
- Neal, D. (1999). The complexity of job mobility among men. *Journal of Labor Economics*, 17(2):237–261.
- Papageorgiou, T. (2012). Learning your comparative advantages. *Working Paper*.
- Pavan, R. (2011). Career choice and wage growth. *Journal of Labor Economics*, 29(3):549–587.

- Poletaev, M. and Robinson, C. (2008). Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys, 1984-2000. *Journal of Labor Economics*, 26(3):387–420.
- Robinson, C. (2010). Occupational mobility, occupation distance and basic skills: evidence from job-based skill measures. *Working Paper*.
- Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135.
- Topel, R. and Ward, M. (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics*, 107(2):445–479.
- Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30(1):1–53.

## A Tables and Figures

Figure 1: Occupations Characterized in Cartesian Coordinates

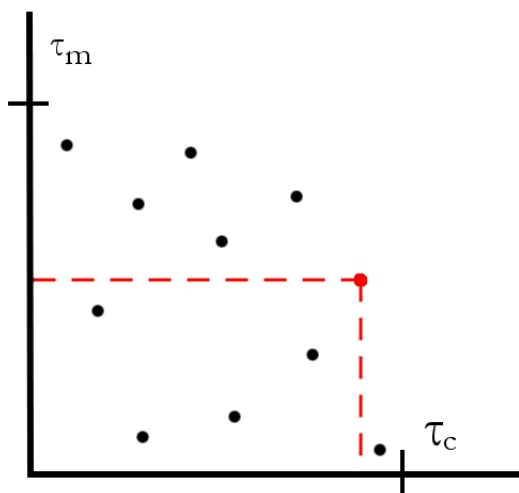
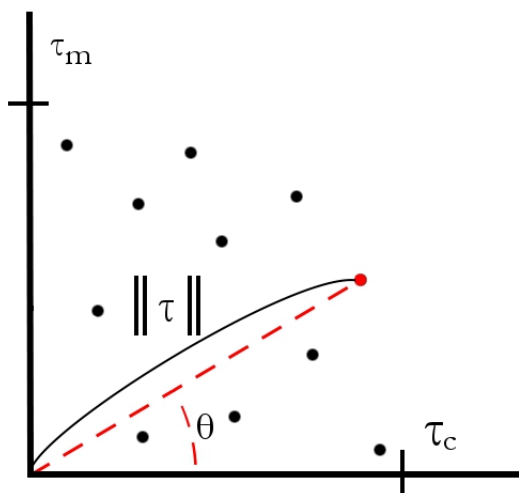


Figure 2: Occupations Characterized in Polar Coordinates



**Table 1: O\*NET Summary Statistics, Occupation Weighted**

	Mean	Std. Dev.	N
Cognitive Tasks, $\tau_c$	.49	.21	440
Manual Tasks, $\tau_m$	.45	.234	440
$\text{Corr}(\tau_c, \tau_m) = -.21$			
The minimum for both task measures is 0 and the maximum is 1.			

Figure 3: Example Question from O\*NET Survey

**5. Mathematics**      **Using mathematics to solve problems.**

**A. How important is MATHEMATICS to the performance of *your current job*?**

Not Important*	Somewhat Important	Important	Very Important	Extremely Important
①	②	③	④	⑤

\* If you marked Not Important, skip LEVEL below and go on to the next skill.

**B. What level of MATHEMATICS is needed to perform *your current job*?**

①	②	③	④	⑤	⑥	⑦
	Count the amount of change to be given to a customer		Calculate the square footage of a new home under construction		Develop a mathematical model to simulate and resolve an engineering problem	Highest Level

Figure 4: Histogram of the Density of Cognitive Skills across Occupations

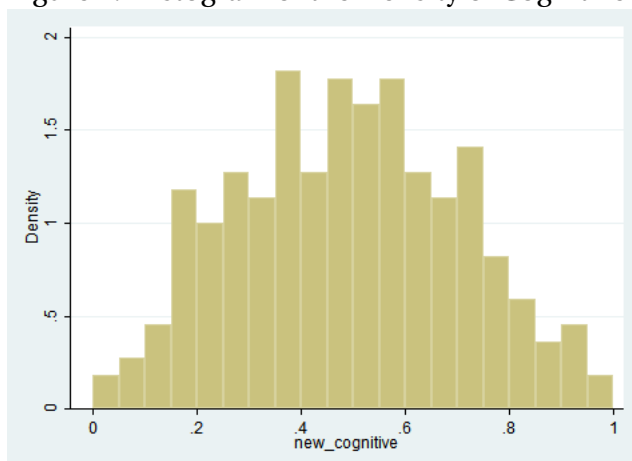
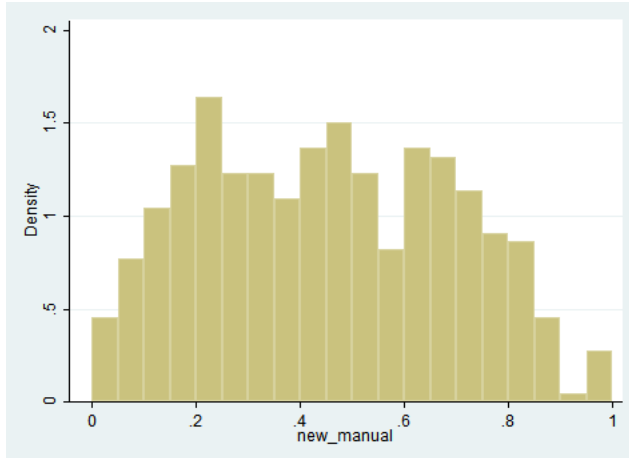


Figure 5: Histogram of the Density of Manual Skills across Occupations

**Table 2: Example Occupations and Constructed O\*NET Scores**

	<b>Cognitive</b>	<b>Manual</b>
Dining Room and Cafeteria Attendants	0.00	0.30
Telemarketers	0.10	0.08
Mechanical Door Repairers	0.33	0.78
Supervisors of Mechanics and Repairers	0.59	0.68
Librarians	0.59	0.26
Economists	0.76	0.02
Hazardous Materials Removal Workers	0.83	0.81
Actuaries	0.94	0.00

**Table 3: Summary Statistics for the NLSY 1979 White Male High School Graduates**

	Mean	Std. Dev.	Min.	Max.	N
AFQT	58	28	0	100	1794
Highest Grade Completed	14	2.4	12	20	1794
Age at Labor Market Entry	20	2.5	15	37	1794
Hourly Wage	10.5	9.4	2.5	70	41740
Cognitive Task Score	.46	.20	0	1	35208
Manual Task Score	.47	.23	0	1	35208
$\theta$	1.2	.27	.19	1.57	35208
$\ \tau\ $	.7	.17	.14	1.26	35208

See the text for additional sample selection criteria.

**Table 4: Measures of  $\theta$  and  $\|\tau\|$  for Example Occupations**

	$\theta$	$\ \tau\ $
Dining Room and Cafeteria Attendants	1.57	0.30
Telemarketers	0.67	0.13
Mechanical Door Repairers	1.17	0.84
Supervisors of Mechanics and Repairers	0.85	0.90
Librarians	0.42	0.64
Economists	0.03	0.72
Hazardous Materials Removal Workers	0.77	1.16
Actuaries	0.00	0.94



**Table 5: Pattern of Changes that Satisfy “Complex” then “Simple” Moves**

# of Total Changes	% that Satisfy C $\Rightarrow$ S	% under random movement
2	87	75
3	60	50
4	46	31
5	31	18.5
6	17	11
7	13	6
8	13	4
9	4	2
10	5	1
11	7	.1
12	4	.01

**Table 6: Fixed Effects Regression for Pr(*ComplexChange*)**

Dependent Variable is Pr( <i>ComplexChange</i> )			
Variable	Coefficient	Std. Err.	p-val
$\sum_{k=0}^{t-1} SimpleChange_{ik}$	-.04	.0017	.000
Time Since Entry	.0026	.0013	.055
Time Since Entry <sup>2</sup>	-.00025	.0000449	.000
Constant	.70	.008	.000
$N_{total} = 30957$	$\sigma_i = .167$	$R^2$ (between) = .43	$R^2$ (within) = .12
$N_i = 1781$	$\sigma_\varepsilon = .45$		$R^2$ (overall) = .18
$\sigma_i$ is the estimated standard deviation of the distribution of fixed effects.			
$\sigma_\varepsilon$ is the estimated standard deviation of the error term.			

<b>Table 7: Fixed Effects Regression for <math> \Delta\theta </math></b>			
Dependent Variable is $ \Delta\theta $			
Variable	Coefficient	Std. Err.	p-val
$\sum_{k=0}^{t-1}  \Delta\theta _{ik}$	-.06	.0019	.000
Time Since Entry	.005	.0003	.000
Time Since Entry <sup>2</sup>	-.00014	.000011	.000
Constant	.07	.0019	.000
<hr/>			
$N_{total} = 30957$	$\sigma_i = .067$	$R^2$ (between)= .52	$R^2$ (within)= .05
$N_i = 1781$	$\sigma_\varepsilon = .11$		$R^2$ (overall)= .002
<p><math>\sigma_i</math> is the estimated standard deviation of the distribution of fixed effects.  <math>\sigma_\varepsilon</math> is the estimated standard deviation of the error term.</p>			

<b>Table 8: Indirect Inference Moments, Part 1</b>				
Moment	Year 1	5	10	20
Mean Log Wages	1.66	2.07	2.46	2.96
Std. Dev. of Log Wages	0.45	0.49	0.53	0.60
Mean Cognitive Task Intensity	0.36	0.43	0.48	0.50
Mean Manual Task Intensity	0.50	0.49	0.48	0.45
Std. Dev. Cognitive Task Intensity	0.19	0.19	0.20	0.20
Std. Dev. Manual Task Intensity	0.20	0.22	0.23	0.23
Mean $\theta$	0.84	0.80	0.75	0.74
Mean $\ \tau\ $	0.65	0.69	0.73	0.71
Std. Dev. $\theta$	0.14	0.16	0.18	0.17
Std. Dev. $\ \tau\ $	0.18	0.17	0.16	0.17

<b>Table 9: Indirect Inference Moments, Part 2</b>				
Dependent Variable is Log Wages				
<b>Coefficient</b>	<b>Estimate</b>	<b>Std. Err.</b>	<b>t-stat</b>	<b>p-val</b>
$\theta$	-.84	.01	-79.51	.000
$\ \tau\ $	.04	.016	2.47	.000
Time Since Entry	0.09	.001	76.53	.000
Time Since Entry <sup>2</sup>	-.00143	.00004	-34.45	.000
$N = 34281, R^2 = .55$				
The constant was estimated but is not reported since it is not used as a moment.				

<b>Table 10: Indirect Inference Moments, Part 3</b>				
	<b>Year 1</b>	<b>5</b>	<b>10</b>	<b>21</b>
Mean of $\Delta\theta$	-0.01	-0.01	-0.01	.0004
Std. Dev of $\Delta\theta$	.142	.146	.141	.126
Mean of $ \Delta\theta $	.08	.08	.07	.08
Median of $ \Delta\theta $	.03	.025	.007	.006
Std. Dev of $ \Delta\theta $	.12	.12	.12	.13
The sample standard deviation of $\theta$ is .27				

<b>Table 11: Simulation of the Effects of Learning on Var <math>\Delta\theta</math></b>				
	<b>Mean Wages</b>	<b>Mean <math>\theta</math></b>	<b>Var <math>\theta</math></b>	<b>Var <math>\Delta\theta</math></b>
Skill Accumulation Only	2.2	.78	.067	.0110
Learning Only	2.2	.78	.082	.0046

**Table 12: Data Moments and Model Moments, Part 1**

	Data				Model			
	Year 1	5	10	20	1	5	10	20
Mean Wages	1.66	2.07	2.46	2.96	1.66	2.10	2.42	2.92
Std. Dev. Wages	0.45	0.49	0.53	0.60	0.39	0.47	0.53	0.65
Mean Cognitive	0.36	0.43	0.48	0.50	0.33	0.46	0.53	0.65
Mean Manual	0.50	0.49	0.48	0.45	0.55	0.55	0.55	0.54
Std. Dev. Cognitive	0.19	0.19	0.20	0.20	0.02	0.07	0.08	0.10
Std. Dev. Manual	0.20	0.22	0.23	0.23	0.08	0.08	0.08	0.08
Mean $\theta$	0.84	0.80	0.75	0.74	1.02	0.87	0.80	0.70
Mean $\ c\ $	0.65	0.69	0.73	0.71	0.64	0.72	0.77	0.85
Std. Dev. $\theta$	0.14	0.16	0.18	0.17	0.06	0.10	0.09	0.10
Std. Dev. $\ c\ $	0.18	0.17	0.16	0.17	0.07	0.08	0.09	0.10

**Table 13: Data Moments and Model Moments, Part 2**

Regression of Log Wage on Time, Time <sup>2</sup> , $\theta$ , and $\ \tau\ $		
Variable	Coefficient, Data	Coefficient, Model
Time Since Entry	0.089	0.010
Time <sup>2</sup>	-0.0014	0.0019
$\theta$	-0.84	-0.82
$\ \tau\ $	0.04	3.11
Standard Deviation of $\Delta\theta$		
	Data	Model
Std. Dev. $\Delta\theta$	0.14	0.07

**Table 14: Data Moments and Model Moments, Part 3**

Regression of $ \Delta\theta _{it}$ on Time, Time <sup>2</sup> , and $\sum_{k=0}^{t-1}  \Delta\theta _{ik}$		
Variable	Data	Model
$\sum_{k=0}^{t-1}  \Delta\theta _{ik}$	-.06	-.14
Time Since Entry	.005	.0025
Time Since Entry <sup>2</sup>	-.00014	-.00006
Constant	.07	.04

<b>Table 15: Parameter Estimates</b>		
<b>Parameter</b>	<b>Meaning</b>	<b>Value</b>
$(\mu_c, \sigma_c)$	Initial Distribution of Cognitive Skills	(3.03, .76)
$(\mu_m, \sigma_m)$	Initial Distribution of Manual Skills	(1.43, .55)
$(R_c, R_m)$	Skill Accumulation Rates	(1.3, .03)
$\sigma_v$	Initial Information About Skills	3.2
$\sigma_\varepsilon$	Noise in the Learning Process	1.54
<b>Wages</b>		
$W = a_0 - a_1\tau_c^2 - a_2^2\tau_m + a_3\tau_c\mu_C + a_4\tau_c\mu_M$		
$a_1$		.06
$a_2$		.005
$a_3$		3.8
$a_4$		1.8

<b>Table 16: Counterfactual Simulations</b>			
<b>Moment</b>	<b>Baseline Case</b>	<b>No Skill Uncertainty</b>	<b>% Change</b>
Mean of $ \Delta\theta $	.032	.019	-41%
Median of $ \Delta\theta $	.03	.021	-30%
Std. Dev of $ \Delta\theta $	.07	.054	-23%