Skill Accumulation, Skill Uncertainty, and Occupational Choice

Carl Sanders*

August 30, 2014

Abstract

Occupational transitions are often viewed either as moves up a job ladder, driven by skill accumulation, or moves across different ladders, driven by skill uncertainty. To quantify the relative importance of these two forces, I create a structural model of multi-dimensional skill accumulation and skill uncertainty. To estimate it, I link data on occupational choices and wages with data on the intensity of cognitive and manual tasks performed in occupations. Identification of the two forces separately is based on the rates at which workers move to occupations that are similar in task mix to their previous occupations. Counterfactual simulations indicate that eliminating skill accumulation would reduce lifetime earnings by 20%, while moving from certainty to uncertainty reduces earnings by only 1%. The small effect of skill uncertainty on earnings is due to workers’ ability to self-insure by occupational choices; if this avenue were shut off the losses from uncertainty would be 2.5%.

JEL Codes: J24, J62, D83

1 Introduction

The occupational transitions of young workers are often viewed as different steps up a single “career ladder.” It is easy to find workers climbing an occupational ladder in labor market micro-data: for example, an individual who moves from being a “Cafeteria Worker” to a “Cook” to a “Food Service Manager.” One natural theory to explain this type of career path is skill accumulation. The skill accumulation model posits that the worker developed skills with work experience and upgraded his occupation as his skills grew.

It is also easy to find careers that do not conform to the career ladder story. Occupational paths where a worker moves from being an “Engine Assembler” to a “Computer Operator” to

*Washington University in St. Louis and the Federal Reserve Bank of St. Louis.
Email: carlsanders@wustl.edu.

Thanks to George-Levi Gayle, John Kennan, Rebecca Lessem, Salvador Navarro, Nicolas Roys, Elizabeth Schechter, Christopher Taber, and many seminar participants for their help and suggestions.

1Individual #39 in the NLSY79.
a “House Cleaner” to a “Manufacturing Inspector” are also common. Careers where workers move between occupations where they perform very different tasks are hard to explain by skill accumulation: what skills does a computer operator develop to help him be a better house cleaner? Rather than skill accumulation, a common explanation for moves between occupations where the required tasks are quite different is that the worker is uncertain about his skills. Experimentation across a variety of occupations is how the worker ensures he does not end up stuck in a job that is not the best fit for his skills.

These two economic forces, skill accumulation and skill uncertainty, can each explain one type of worker mobility: skill accumulation for movement up career ladders, and skill uncertainty for movement across ladders. The labor economics literature has found that each of these forces analyzed in isolation can explain overall career patterns of occupational choice and wage growth. But by evaluating the two forces separately, we do not know which of the two is actually more important for worker occupational choices and wage outcomes. In this paper, I quantify the relative contributions of the two forces to workers’ occupational transitions and wage growth over their whole career.

To do this, I construct and estimate a life-cycle model of worker occupational choice that includes both skill accumulation and skill uncertainty. Simulation of the model allows two thought experiments to distinguish the two forces: if a labor market entrant knew his true skills, how would his lifetime occupational transitions and wages change? Conversely, how would his career change if his skills did not increase but he still had to learn about his true skill levels? In the model, workers have a multi-dimensional skill bundle and are able to both invest in those skills and learn their skill levels by choosing between occupations that require different amounts of tasks performed. The model extends and combines some of the most recent work on task-based human capital accumulation from Yamaguchi (2012) with a model of worker learning about skills from Antonovics & Golan (2012).

The model is estimated by linking information on worker wages and occupational choices from the National Longitudinal Survey of Youth 1979 with data on the intensity of cognitive and manual tasks performed in occupations from the O*NET database. After estimating the model parameters, in particular the rates of skill accumulation and the initial level of uncertainty about skills, I simulate the model twice, first eliminating skill accumulation and second skill uncertainty. The results imply large welfare losses from fixing worker skills, around 20% of lifetime earnings. On the other hand, the estimated welfare losses from uncertainty are much

\(^2\text{Individual \#2853 in the NLSY79.}\)
smaller: moving from perfect certainty to uncertainty reduces lifetime earnings by only 1%. This does not mean uncertainty is irrelevant to workers: in response to skill uncertainty, they experiment through occupational choices to self-insure against investing in the wrong skills. In fact, if workers were prevented from changing their occupational choices in response to uncertainty, the losses from uncertainty would increase to 3.5% of lifetime earnings.

This paper makes two significant contributions to the learning and human capital literatures. First, this is the first paper to determine the relative importance of skill accumulation versus skill uncertainty where both forces are given an equal chance to explain occupational transitions. Almost universally, empirical human capital papers assume full information and empirical learning papers do not allow for human capital growth. The only partial exceptions are learning papers by Pavan (2011) and James (2011) which allow for firm-specific and general human capital growth. Ostensibly, this means they contain both learning and human capital accumulation. But their primary focus is learning, and the models of human capital they use are too simple to allow for both forces to explain observed occupational choices. Unlike those models, the task-specific model of human capital I use can explain wage growth patterns and occupational mobility measures simultaneously.

For my second primary contribution, I use a new identification strategy based on occupational tasks to separate the two forces. First, I separate the “magnitude” of an occupation’s tasks, the total intensity of tasks performed there, from the “specialization”, the relative intensity of its different tasks. For example, consider professional athletes versus economists. Both occupations require performing high-level tasks and so are both high magnitude occupations. But they are intensive at different things: professional athletes are specialized in intensive manual tasks, and economists are specialized in intensive cognitive tasks.

I then use the definitions of magnitude and specialization to define “return moves”: when a worker makes two occupational moves, the first move with an increase (decrease) in cognitive specialization and the second with a decrease (increase) in cognitive specialization. The focus on specialization instead of magnitude reflects that skill accumulation will lead workers to perform higher-level tasks over their career, but should not dramatically change the relative intensities of those tasks over time. Skill uncertainty, in contrast, will have workers experimenting with very different tasks than the ones they believe are long-run optimal to ensure they are not wrong, so they will make a large number of return moves when those experiments do not work out. I use the prevalence of return moves from the data in model estimation to identify the skill uncertainty force. This identification method is an innovation to the “return moves”
identification strategy used in some previous papers on learning (e.g. Crawford & Shum, 2005).

Although the results show that skill accumulation is the primary driver of wage growth over the career, they also emphasize the importance of occupational mobility as a way workers can self-insure against learning they are low-skilled. Since the overall welfare loss of skill uncertainty seems small relative to the welfare loss from eliminating skill growth, it would be tempting to assume skill uncertainty can be ignored. A more careful interpretation, however, is that if workers are free to change occupations, they can hedge against most of the losses they would incur from skill uncertainty. This suggests an additional, potentially significant cost of labor market frictions and restrictions: workers who cannot experiment with different occupations can experience welfare losses because they cannot learn their comparative advantage as quickly as they would like.

2 Literature Review

Learning models were developed to analyze worker transitions between firms, and came to also consider worker transitions between occupations, industries, and many other categories. For Jovanovic (1979) the focus was on workers and firms learning about the quality of their match, while Miller (1984) looked at workers who were uncertain about their match with occupations. Gibbons & Waldman (1999) and Papageorgiou (2013) considered workers trying to find the right tasks to perform within a firm given an unknown skill set. Lluis (2005) empirically implemented the Gibbons & Waldman (1999) model of within-firm mobility, but naturally could not address inter-firm or inter-occupation mobility. Neal (1999) instead focused on “careers” and derived and tested predictions from a model of workers choosing between uncertain career paths. Pavon (2011) estimated a structural version of this model and quantified the relative importance of career versus firm transitions. James (2011) extended the Miller (1984) framework into a realistic model of workers learning about their optimal occupation. My paper is most closely related to Antonovics & Golan (2012), who model workers as having uncertainty about a low-dimensional set of skills, and performing different types of tasks that can be more or less informative about those skills.

The human capital literature has primarily focused on wage growth. In particular, what skills transfer between jobs, and how can growth in these skills explain growing wages over the career? The original Becker (1964) model posited that workers have a stock of general human capital useful at all firms and another set of stocks that are each only valued at one specific firm.
Since that initial distinction, labor economists have suggested that workers have human capital that may be industry specific (Neal, 1995), occupation-specific (Kambourov & Manovskii, 2009), and most recently task-specific (Gathmann & Schoenberg, 2010 and Yamaguchi, 2012). In task-specific (sometimes also called skill-specific) human capital models, workers have distinct categories of skills, such as cognitive skills, manual skills, or interpersonal skills, and different jobs produce output using those skills in different amounts. Sanders & Taber (2012) reviewed the human capital literature and argued that task-specific human capital can nest the previous models and offers the best way to unify disparate empirical results.

While both learning and human capital have active empirical literatures, there is basically no research attempting to quantify the relative importance the two forces. Nagypal (2007) separated learning about the firm-worker match from firm-specific human capital as an explanation for the empirical firm tenure distribution. Some learning papers have allowed for simple versions of human capital growth, in particular Pavan (2011) and James (2011). They identify the importance of uncertainty by looking at worker mobility rates across firms, careers, or occupations, and with high observed levels of worker mobility the model estimates learning to be very important. But their simple human capital models were not meant to explain transitions: general human capital cannot explain transitions at all, and firm-specific human capital can only slow down transitions because workers need to pay an additional opportunity cost to leave their current firm. As opposed to these older models of human capital, the task-specific skill accumulation model I use here can explain both worker wage growth and worker mobility within one model. The fact both the task-specific human capital and skill-based learning models used in this paper can fully explain the cross-sectional patterns in my data allows for a fair comparison of the two.

3 A Simple Model of Skill Accumulation and Skill Uncertainty

To explain the model mechanisms and data requirements, it is useful to look at a simplified version of the estimated model. The simple version will contain all the key elements of the estimated model: workers with an unknown skill bundle choose occupations of different task intensities to solve a life-cycle wage maximization problem. Every period, workers choose an occupation based on their beliefs about their skills, and between periods those beliefs transition

---

3Another explanation for many of the same trends discussed here is labor market search, which is not modeled here. Bowlus & Liu (2012) quantify at the relative contributions of search and human capital accumulation to wage growth and so can be viewed as a complement to this paper.
depending on the chosen occupation. The transition has two components: skill accumulation and Bayesian learning about uncertain skills. Workers choosing higher task intensity occupations have two advantages: First, they have a faster rate of deterministic skill accumulation. Second, they receive more precise signals from which they can learn their true skills.

The simple model in this section has a one-dimensional human capital framework, two occupations, two periods, and simplified distributions and restrictive functional forms, but the intuition gained by analyzing this simple version carries over into the estimated multi-dimensional model. In a later section I describe the exact parametrization used in estimation, but of particular note is that there I allow for a two-dimensional skill vector and a continuum of occupational choices.

Workers have a skill bundle consisting of a known part \( H \in \mathbb{R}_+ \) and an unknown part \( S \in [0, 1] \). The worker has beliefs \( B \) about the unknown part which can be summarized by the Beta distribution over the \([0, 1]\) interval with parameters \( \alpha \geq 0 \) and \( \beta \geq 0 \). With these priors the workers have expected skills \( E_B[S] = \frac{\alpha}{\alpha + \beta} \).

Workers can choose either a high or low task intensity occupation in each period, denoted \( \tau = 1 \) or \( \tau = 0 \) respectively. Wages as a function of worker beliefs, known skills, and chosen occupation are given by

\[
W(B, H, \tau) = \tau (E_B[S] + H) - \tau.
\]

Wages in the low task sector are 0 for all workers no matter their skills, but high task sector wages are \((E_B[S] + H) - 1\), which may be positive for some workers and negative for others. This wage equation leads to positive assortative matching in the static problem: all the low expected-skill workers choose the low task sector and all those with high expected skills choose the high task sector.

There are two primary arguments for a wage function of this form: first, it predicts that higher skilled workers take higher task jobs and earn higher wages than low skilled workers. While there are no skill measurements over the career in the data, Table 6 documents that individuals with higher cognitive ability measured before labor market entry end up in occupations with higher level cognitive tasks and earn more. Second, Yamaguchi (2012) finds that a similar specification works well in fitting life cycle profiles of wages and task choices, although he makes more assumptions on other functional forms and the worker choice problem. The estimated wage equation will be a generalized version of equation 1.
The second-period wage function given optimal sector choice is

\[
W^* (B, H) = \begin{cases} 
0 & \text{if } E_B [S] + H < 1 \\
E_B [S] + H - 1 & \text{if } E_B [S] + H \geq 1 
\end{cases}.
\] (2)

In the first of the two periods (with no discounting), the worker’s value function is

\[
V_1 (B_1, H_1) = \max_{\tau_1 \in \{0, 1\}} \tau_1 (E_{B_1} [S] + H) - \tau_1 + E_{B_1} \left[ W^* (B_2 (X, B_1), H_2) \right]
\] (3)

where the terms in the continuation value require explanation. The known part of worker skills transitions according to

\[
H_2 = \begin{cases} 
H_1 + R & \text{if } \tau_1 = 1 \\
H_1 & \text{if } \tau_1 = 0 
\end{cases}.
\] (4)

where \( R \) is a fixed parameter known to the worker. The human capital transition process is learning-by-doing: workers who perform no tasks (\( \tau = 0 \)) gain no skills but those who perform tasks increase their skills. Positive \( R \) means the worker may be willing to forgo wages today by choosing the high task sector because it will increase wages tomorrow enough to compensate. The estimated model will contain a more flexible specification of learning-by-doing as well as exogenous skill growth.

The transition rule for \( B \), the beliefs about the unknown skill, is the learning component. The worker receives a signal between periods that depends on his sectoral choice. If the worker chooses the low task sector, he receives no information and \( B_2 = B_1 \). If he chooses the high task sector he can see an informative signal \( X \in \{0, 1\} \):

\[
X \sim \text{Bernoulli} (S).
\] (5)

The signal can be interpreted as the worker taking an exam where he only sees the pass/fail result, and he passes with a probability given by his unknown skills. He then uses the result to update his beliefs. Bayesian updating rules for this model give a closed form for the transition of beliefs:

\[
S_2 | X \sim \text{Beta} (\alpha + X, \beta + 1 - X)
\] (6)
which can be interpreted more easily using the averages conditional on test result:

\[ E_{B_2|X \{S_2\}|X} = \begin{cases} \frac{a+1}{a+\beta+1} & \text{if } X = 1 \\ \frac{a}{a+\beta+1} & \text{if } X = 0 \end{cases} \quad (7) \]

Success will lead to upwards revision of beliefs, and failure will lead to downwards revision. This learning process is a simplified version of Antonovics & Golan (2012); in the estimated model workers have normally distributed beliefs and signals. The interpretation will be exactly the same: workers who perform tasks more intensely get more opportunities to see success or failure, and so gain more information. To my knowledge there is no direct evidence on the informativeness of performing tasks for worker skills, but the assumption of increasing information as a function of tasks seems reasonable and the estimated model can fit the data.

When choosing his sector the worker does not know the value of \( X \) and it can be shown that unconditional on \( X \),

\[ E_{B_1 \{S_2\}} = E_{B_1 \{S_1\}} \quad (8) \]

in either sector. In expectation the worker does not expect his beliefs to change over time no matter what sector is chosen. It will then require some additional work to show the incentives for occupational experimentation. For now let both the known skill level and the accumulation parameters be 0, so \( H_1 = H_2 = R = 0 \) in the above formulas. Writing out the full period 1 value function,

\[ V_1(B_1) = \max_{\tau_1 \in \{0,1\}} W(B_1,0,\tau_1) + \tau_1 \left[ \Pr(X = 1) \times W^* \left( B_2^h \right) + \Pr(X = 0) \times W^* \left( B_2^l \right) \right] + (1 - \tau_1) W^* (B_1) \quad (9) \]

where \( B_2^h \) ("high") is the updated beliefs after a successful signal and \( B_2^l \) ("low") is the update after a failed signal.

Now consider a worker indifferent between the two sectors in terms of wages, so that \( E[S] = 1 \) and he would earn 0 in either sector. If the worker chooses the uninformative sector, his beliefs will not change so he will receive 0 wages again in the final period and have total value of 0 between both periods. If he chooses the informative sector, he takes the exam. If he fails, he can still guarantee himself 0 in the second period by taking the low-task job. On the other hand, passing the exam increases his mean belief and will move him into the high-task sector where he receives a positive wage and the continuation value for the high-task sector is strictly positive.
This means this worker who would receive a wage of 0 in both sectors in the first period strictly prefers the high-task sector when this learning motive is introduced.

The value of learning comes from the existence of a “safe” occupation. Taking the test by choosing the $\tau = 1$ sector is a risky bet compared to $\tau = 0$: with $\tau = 0$ he receives no updates and can calculate his wage with certainty in the second period. In choosing the high-task sector, a worker will revise his beliefs either upwards or downwards. Revising downwards has limited downside risk because he can always guarantee himself a wage of 0. On the other hand, there is no equivalent upside limitation: wages above the threshold are linear in skills without bound. The existence of a safe job that does not depend on the worker’s skills creates an incentive for experimentation: the worker will be willing to forgo current wages in order to expose himself to this limited downside uncertainty.

Reintroducing the skill accumulation component alongside skill uncertainty does not change the direction of the incentives of either. But they do not act independently of each other. Take a worker who is extremely low skilled so that no matter whether he passes or fails the exam he would still be in the low-task sector in the second period. Without skill growth, the value of the learning motive is 0 for this worker. Now consider increasing the skill accumulation parameter so that if the worker was in the high-task sector in the first period, his skills would increase to where he was indifferent between sectors in the second period. Then the learning motive would bite: there would only be upside risk since now passing the test would lead to positive wages but failing he would remain at 0. This general intuition for the relationship between human capital accumulation and learning holds in the more general model: some workers can have the joint incentives of skill accumulation and skill uncertainty be stronger together then the sum of the two forces separately.

Previous models have only looked at the two factors in isolation. This toy model along with results from the general model show that both forces analyzed individually have similar quantitative predictions for worker occupational mobility and wage growth: in both cases, the incentive is for workers to take high-task jobs in order to increase their future wages. This investment motive then leads to wage growth between periods. Since both forces individually will be found to be able to explain the cross-sectional patterns of wages and occupational choices, distinguishing the forces is going to require more than simply matching wage growth and occupational transition rates.
4 Data

4.1 Construction of the Estimation Sample

Standard data sources can provide information on wages and occupational choices; however, no data set contains workers’ beliefs about their skills or updates to those beliefs over their careers. Instead of estimating the model using beliefs, I use the implied relationship from the model between the chosen occupations and workers’ beliefs about their skills. This requires lifetime information on workers’ wages and the tasks performed in their occupations.

To get this data, I merge the National Longitudinal Survey of Youth 1979 (NLSY) with the U.S. Department of Labor Occupational Characteristics Database, O*NET. The NLSY allows me to observe the occupational choices and wages of workers over their whole careers. However, the NLSY does not have records of tasks performed by workers, but only records of the “three-digit” US Census classification of occupations, the most disaggregated occupational classification used in the US. These occupational classifications range from detailed categories such as “Automotive Glass Installers and Repairers” to general ones such as “Retail Salespersons.”

To complement the occupation data, the O*NET database contains data on a variety of tasks performed in US occupations. I take task data from O*NET for each occupation and match it with the three-digit occupations in the worker histories from the NLSY. The result is the work histories and demographic data from the NLSY with not just wages and occupation categories, but also measures of tasks performed over time. Data selection and merging the data sets required dealing with a large number of details. In this section, I outline the details of the O*NET data and how I construct the final sample of worker careers.

The O*NET database is a representative survey of occupations in the United States by the U.S. Department of Labor. The data is constantly being updated; currently it contains results from surveys administered between 2007 and 2014. A nationwide random sample of workers completed a survey consisting of four different questionnaires asking about their job characteristics, e.g. “Education and Training” or “Work Activities.” All means and variances from both stages are released to the public, but not the individual-level responses. The end result is data at the occupation level that contains detailed information on the tasks and activities workers perform in those occupations.

The publicly released data set has hundreds of scores for each occupation with questions ranging from “How often do you spend time sitting?” to “Does your job require creativity?” An example question taken from the “Work Activities” questionnaire is shown in Figure 1. To
reduce this large number of measurements of occupational characteristics to a manageable
number, I follow the literature, e.g. Yamaguchi (2012), and reduce these questions to a two-
dimensional measure of “cognitive” tasks and “manual” tasks performed. I take these as cor-
responding to measures of the task intensities $\tau$ in the model and normalize them between 0 and
1.

For worker-level career data, I use the National Longitudinal Survey of Youth 1979. 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, with the final survey
I use being 2008. Each interview, the respondents were asked to recall details of their weekly
labor market history since the previous interview. This Weekly Labor Status variable can be
used to create a weekly history of workers labor force participation status, when they changed
employers, and some characteristics of the job such as the hourly wage, hours worked, occu-
pation, and industry. Additionally, the NLSY contains demographic variables such as age, sex,
race, education, marital status, parental characteristics, and scores on the Armed Services Vo-
cational Aptitude Battery tests, a battery of exams on a number of different topics. I take a
standard sample of white males with a high school degree and no college attendance who have
left school and have a firm attachment to the labor force; see Appendix A for additional sam-
ple details. Restriction to high school graduates is to avoid the issue of the amount of learning
that may take place during college. There is evidence that students learn a significant amount
about their own ability in college, see for example Arcidiacono et al. (2010) and Stinebrickner &
Stinebrickner (2012).

From the NLSY I create a data set of monthly observations that contain the average wage
paid during the previous month, the occupation and employer the individual worked for the
most that month, the actual labor force experience of the worker, and the time since labor mar-
ket entry. I link the task scores constructed from the O*NET at the occupation level with the
occupations in the NLSY to construct the data used for estimation. Descriptive statistics for
this sample will be summarized in Section 5. However, before that it is necessary to address the
way that this linked data induces measurement error.

4.2 Measurement Error

Use of the linked O*NET-NLSY leads to three important measurement error issues. First, the

---

4Details are in Appendix A. The most natural next category of skills would be interpersonal skills, but these are difficult to find in the data and remain an area for future research.
5With a separate oversample of disadvantaged minority groups, which I do not use.
NLSY itself has issues with accurate recording of occupational transitions. Second, the O*NET task measures are at the occupation rather than the individual level. Third, the two data sources use significantly different occupational coding schemes and merging across them may induce errors. These problems are common in the existing literature, e.g. in Yamaguchi (2012) or Gathmann & Schoenberg (2010), but it is necessary to discuss these in order to understand my results.

First, the rate of occupational transitions may be overstated. The NLSY records occupations by asking the respondent to describe in words the tasks they perform and recording responses verbatim. In a later coding stage, occupational experts coded those responses into three-digit occupational categories. However, the coders do not use previous responses from the same respondent when classifying, so in theory the same verbatim description of tasks could be scored differently across years as occupational categories are not always clearly distinguished. This could lead to spurious recorded occupational transitions and biasing upwards the measured rate of occupational transitions. To deal with this, I eliminate all within-firm occupational transitions that are of the form “ABA:” a worker starts in occupation A, moves to B, and then back to A. These types of moves are actually consistent with my model and could be evidence for the importance of skill uncertainty. However, Neal (1999) and Kambourov & Manovskii (2009) argue that these transitions are best understood as spurious. The net effect of removing these types of changes in my setting is to reduce the estimated importance of skill uncertainty.

The second form of measurement error is noise in the task measures themselves. As in almost all studies using task data, tasks are measured at the occupational level and then merged into individual-level data. But there is a significant amount of task heterogeneity within occupations, recorded both by Autor & Handel (2009) in a specialized data set as well as in the O*NET itself by measures of the variance of responses. Here I follow the literature and take the average task measure as the true measure of the task intensity of the occupation for all workers. Without additional data on worker-level task heterogeneity it does not seem possible to generalize this assumption.

The last form of additional measurement errors are inconsistencies when merging across different occupational coding systems. There are three different occupational coding schemes used between the O*NET and NLSY; merging details are in Appendix A. There is no unique way to map an occupation from one scheme to another. For example, in the Census 1970 scheme “Editors and Reporters” is one occupation, but in the Census 2000 revision this was split into two different occupations. This could lead to either spurious occupational moves or
unrecorded true moves. However, most occupations have natural matches across coding systems. Additionally, for robustness I also did the data analysis while dropping any matches that seemed less plausible and it did not change any of the results.

5 Descriptive Statistics: Specialization and Magnitude

To quantify worker transitions across occupations, I use a new characterization of occupations in order to determine when workers make “big” moves across occupations versus “small” moves. From the data, we know when a worker’s occupational title changes, but this title is just an *ad hoc* name: we would like to know how much the worker changed his actual production behavior.

One common approach in the labor literature is to use some level of aggregation; for example, moves from white collar to blue collar are assumed to be significant whereas within-category moves are insignificant, or moves across one-digit categories versus within those categories. With smaller panel data sets such as the NLSY, the three-digit occupation cells can have very small numbers of workers in them, and many are not represented at all, so statistical power requires some aggregation. Aggregation is also useful in dynamic programming models with occupation-specific human capital, such as James (2011), where allowing for the over 400 three-digit occupations to have their own stock of human capital is computationally impossible.

While there has been no detailed analysis comparing different aggregation methods, Figures 2 and 3 are evidence that task measures can potentially offer a more accurate way of comparing occupations. In Figure 2, the cognitive and manual task scores of all three-digit occupations from O*NET are plotted with a marker for whether they are blue collar or white collar. Notably, there is vastly more variation in tasks within those categories than across. Figure 3 shows that even one-digit occupational categories have a significant amount of within-category variation in cognitive and manual tasks.

The task approach quantifies the economic intuition that what the worker is actually producing matters more than the title when thinking about “similar” jobs. Occupations are then compared by some metric on the underlying similarities or differences in the tasks the average worker performs in those jobs. For example, Gathmann & Schoenberg (2010) use the Euclidean distance in the task space to characterize the size of a move. But both skill accumulation and

---

6A US Census aggregation of occupations at a coarser level than three-digit occupation.
skill uncertainty can explain significant moves under this measure, since both gaining skills rapidly or getting new information about your skills can lead to significant changes in your optimal tasks.

Instead of the Euclidean distance, then, I use a new parametrization of the task vectors that will help to identify skill accumulation versus skill uncertainty. First, I define the task “magnitude” of an occupation: that is, the total task intensity, ignoring how it is split across different tasks. For example, consider two occupations, Job C where workers perform 1 unit of cognitive tasks and no manual tasks, and Job M, where they perform no cognitive tasks and 1 unit of manual tasks. For both occupations the task magnitude is 1. Then, I also calculate the “specialization” of the occupations: the share of the task vector that is performing manual tasks. So in the example, the specialization of Job C is 0 and Job M is 1. This two-dimensional score captures the idea that workers at both Job C and Job M are performing hard tasks, but they are doing very different things. An extreme example is comparing professional athletes and economists. Both are at the highest end of the distribution in terms of total worker skills, but their jobs have almost no task overlap. The identification argument in the next section will make use of the magnitude/specialization decomposition in particular.

The quantitative measures I use for magnitude and specialization are simply changing each occupation score from the Cartesian coordinates \( \tau = (\tau_C, \tau_M) \) into the polar coordinates \( \tau = (\theta, \|\tau\|) \). The transformation is given by

\[
\cos \theta = \frac{\tau_C}{\|\tau\|}, \quad \|\tau\| = \sqrt{\tau_C^2 + \tau_M^2}.
\]

This is very close to decomposing each occupation into the ratio of tasks and the sum of tasks. Directly using \( \frac{\tau_C}{\tau_M} \) as a measure of specialization is unstable as \( \tau_M \) is small for many occupations, but dividing both the top and bottom of the expression for \( \cos \theta \) by \( \tau_M \) shows this measure is just a better-behaved transformation of the ratio. A graphical representation of these new scores is shown in Figure 4.

The magnitude/specialization parametrization is new, so it is useful to understand how it relates to the usual methods of aggregating occupations. Table 6 summarizes worker outcomes at 10 years after labor market entry by one-digit occupation. The one-digit classifications vary dramatically in terms of wages, tasks performed, average firm tenure, and the cognitive test scores of their workers. For example, the most cognitively specialized occupation is sales, with the average worker about 1.5 standard deviations more cognitively specialized than the over-
all average. Notice that even if workers were only “promoted” by moving to higher magnitude occupations but never changed specializations, there could still be one-digit occupational transitions. This is one advantage of the task-specific human capital model: some standard occupational moves that firm- or occupation-specific human capital would be unable to explain seem natural.

Workers make significant moves in task measures across their careers. Table 1 contains summary statistics of workers over time in the NLSY sample, with measures of their average wage levels, wage dispersion, and average levels of the magnitude and specialization measures. Of course, mean wages and the dispersion of wages increase over the career. Less well known, the average magnitude of task intensity increases by about $\frac{1}{2}$ of a cross-sectional standard deviation. This is consistent with the idea that workers move over time to jobs where they perform higher overall levels of tasks. Additionally, there is a slight movement from more manually-specialized tasks towards more cognitively-specialized tasks as workers age.

Another way to understand these measures is to look at their relationship with wages. That is, where in the task space are high and low wage workers? Table 2 answers this in the pooled cross-section with an OLS regression of wages as a function of current period tasks performed. The results show that higher-wage workers are located in higher magnitude jobs and more cognitively specialized jobs, but a standard deviation change in magnitude has about a 7 times stronger predicted change than a standard deviation change in specialization. Again this seems intuitively plausible: workers who do “more” should get paid more, but conditional on that shifting the balance of what they do only has a minor effect.

But one explanation for these cross-sectional patterns is simply sorting: high-wage type workers prefer high-magnitude and cognitively specialized sectors. As a first step in dealing with this problem, Table 3 runs a fixed effects version of this same Mincer regression. While the fixed effects model is still not an estimate of the structural wage equation from the model, the differences between the OLS and FE regressions are informative. In particular, the returns to magnitude fall from $1.41$ in hourly wage for a one standard deviation increase to $0.95$, and the returns from a one standard deviation movement of the specialization measure $\theta$ towards cognitive tasks increases from $0.20$ to $0.27$. The differences between these two regressions indicate that within worker careers, increases in magnitude and cognitive specialization are still associated with wage growth, but some around 30% portion of the measures returns to magnitude is driven by simple selection. Fully dealing with selection into sectors will be a task for the estimated model.
Detailed patterns in worker mobility can also be informative about how workers are making choices across their career. Table 4 looks at the monthly probabilities of making an occupational change. In part (a), the probability of changing three-digit occupation starts at around 5% per month and decreases in the long run to about 2.3%. The first two columns of part (b) of the table show that there is little monthly change in specialization and magnitude, and both types of changes slow down dramatically as workers age. The final column shows that not just the average move, but also dispersion of moves decreases as workers age.

None of these descriptive statistics do much to differentiate skill accumulation from skill uncertainty. As simulation of the model will later show, both can justify the cross-sectional relationship between wages and tasks described here, as well as the decreasing level of mobility across the career. Identification of these two forces separately will rely on another aspect of the data described in the next section.

### 6 Identification

The identification argument proceeds in three steps: first, I argue that in the model with two-dimensional skills and myopic agents, the relative amount of mobility across magnitude versus mobility across specialization identifies the relative importance of skill accumulation and skill uncertainty. Second, endogenizing task choices destroys this clean identification, such that there is no longer a one-to-one mapping between the specialization and magnitude mobility rates and the model parameters. Third, I argue that a new “return moves” measure can identify the relative contributions of these two forces. Skill uncertainty will lead to individuals returning to previous choices while skill accumulation will not. The trick is to define “previous choices” as previous values of the task specialization measure rather than the specific occupation titles.

In the myopic case, workers skills’ and thereby task choices will increase over time. Graphically speaking, workers occupational choices will move out from the origin over time and his measured task magnitude will increase. On the other hand, with myopic skill uncertainty every period the worker learns something new about his skills, so he changes his tasks accordingly. But with only skill uncertainty, since workers have rational expectations about their skills, the average worker’s task magnitude will not increase over time. Separately identifying the two components in this simplified model is then straightforward: the average growth in magnitude gives the skill accumulation component, and the changes in specialization identify the skill uncertainty component. This corresponds exactly to the simplest intuition: the model parameters
are identified by moves up career ladders versus across ladders.

But this simple intuition has two drawbacks: it is not robust to either non-myopic task choices or task measurement error. Consider the simple model where workers consider the dynamic effects of their choices. For some values of the human capital accumulation parameters, workers move downwards in magnitude or move across different specializations. The intuition is that if the human capital accumulation rate is high enough for (say) cognitive skills, early in life it may make sense to specialize in those tasks even if the individual plans to primarily use manual skills in the long run. Similarly, skill uncertainty and experimentation can lead to workers changing their average magnitude: it may be worthwhile for everyone to try very difficult tasks early in the career to gain as much information as possible. This is borne out by the example worker paths (occupational choices for years 0 through 9) seen in Figures 5a and 5b, simulated off of the full model with either only human capital accumulation or only learning.

Measurement error in occupations also causes problems with identification based on moves across task specialization. If there is measurement error in task choices, even the human capital accumulation-only path will have changes in specialization, which this identification strategy would attribute only to learning. In the “Data” section I argued there is good reason to believe that there is significant measurement error of occupational tasks.

For these reasons I do not use the relative proportions in moves up ladders versus across ladders to identify the relative importance of human capital accumulation versus learning. Instead, I create a new measure of “return moves”. Consider a worker who makes an occupational transition from a manually-intensive job to a cognitively-intensive one. Human capital can explain this transition: either through exogenous skill growth or investment, the worker has higher cognitive skills now. But then if the worker’s next move is back towards a manually-specialized job, the human capital explanation becomes difficult. On the other hand, skill uncertainty has a natural explanation for the move back to a manually-specialized job: the worker learned his skills were not the best match for those manual tasks.

Specifically, I define a “return move” to refer to two consecutive moves where the first move is towards a more cognitively specialized job and then the second is back towards a more manually specialized job, or vice versa. As the estimated model will show, human capital moves (even with endogenous accumulation) tend to be in one direction: workers become either more cognitively or more manually specialized throughout their career. On the other hand, we see a large number of return moves in the data: over 70% of second moves early in careers are return moves. This percentage significantly exceeds what either random job movement or measure-
ment error could account for. Random job movement or large amounts of measurement error would predict that at most 50% of moves will be return moves. Estimation of the model will show that such a high proportion of return moves is inconsistent with the skill accumulation only version of the model. In the full model, skill uncertainty will then play the primary role in matching the return moves moment.

7 Full Model

The estimated model extends the simple model in a number of ways. However, the extensions do not change the incentives for skill accumulation and skill uncertainty explained in the simple model.

Time is discrete and runs yearly from $t_0 = 20$ to retirement at given age $T = 65$. At labor market entry, a worker is endowed with with two-dimensional skill bundle $S^* \equiv (C^*, M^*)$ where the values are drawn from independent normal population distributions\(^7\)

$$C^* \sim N\left(\mu^*_C, \frac{1}{p^*_C}\right), \ M^* \sim N\left(\mu^*_M, \frac{1}{p^*_M}\right).$$

His only knowledge of his true starting skills levels, however, comes from one signal for each skill. In particular, before entry each worker observes

$$Z_c \equiv C^* + \epsilon_{C0} \text{ and } Z_m \equiv M^* + \epsilon_{M0}$$

but cannot directly observe the components of the $Z$’s. The $\epsilon_0$ are mean-0 normal random variables with precisions $p_{C0}$ and $p_{M0}$, respectively. This implies that the worker’s initial beliefs about his skills are given by normal random variables $B_{C0}$ and $B_{M0}$ such that

$$E[B_{0C}] = \mu^*_C, \text{ Precision } (B_{0C}) = p^*_C + p_{C0}, \ E[B_{0M}] = \mu^*_M, \text{ Precision } (B_{0M}) = p^*_M + p_{M0}.$$ \hspace{1cm} (13)

In period $t$ the state variables for the worker are the current beliefs about his skill vector, which are summarized by the random variables $B_{Ct}$ and $B_{Mt}$. Beliefs in every period will be normally distributed because of assumptions below. The state vector $B_t$ in period $t$ can be summarized as a 4-dimensional vector $B_t = (\mu_{Ct}, p_{Ct}, \mu_{Mt}, p_{Mt}) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$ where $\mu_{Ct}$ and $\mu_{Mt}$ are the worker’s current expectation of his skills and $p_{Ct}$ and $p_{Mt}$ summarize his precision

\(^7\)Working in terms of precisions rather than variances simplifies the algebra.
of beliefs.

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, \( \tau = (\tau_C, \tau_M) \) 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 positive \((\tau_C, \tau_M)\) is a feasible choice for any worker in any period, so the control vector \( O_t \) (for “occupations”) is \( O_t = (\tau_C t, \tau_M t) \in \mathbb{R}_+^2 \).

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. Wages, tasks, and beliefs are related by

\[
W_t(B_t, O_t) = w(t) + (1 + k_C \hat{\mu}) \times \left( a_1 \hat{\mu} \tau_C - a_2 \tau_C^2 \right) + (1 + k_M \hat{\mu}_M) \times \left( a_3 \hat{\mu}_M \tau_M - a_4 \tau_M^2 \right). \tag{14}
\]

where \( w(t) \) is an exogenous time trend in wages, the \( a, k, w_g \) and \( w_{gr} \) are parameters, and \( \hat{\mu} = (\mu + 3) \cdot 1 \{ \mu > 3 \} \) is a piecewise linear function ensuring the returns to skills are non-negative. This wage equation similar to the quadratic form in Yamaguchi (2012). It is unapologetically reduced form; to my knowledge there is no micro-founded functional form relating tasks, skills, and output. The wage-maximizing task choice for each type is an increasing function of the corresponding skill level. The rest of the parameters and functional form govern the curvature of the policy function and the value function. Since in this model learning is a story of risk-loving behavior (see the explanation in the Section 3 above), the curvature of the wage equations (and thereby the value function) is important to model flexibly.

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

\[
V_t(B_t) = \max_{O_t} W_t(B_t, O_t) + 0.95 \cdot E_t [V_{t+1}(B_{t+1})], \quad t = 1, 2, ..., T - 1 \tag{15}
\]

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

where \( \mathcal{B} \) 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 final period the worker solves the static problem

\[
V_T(B_T) = \max_{O_T} W_T(B_T, O_T). \tag{16}
\]
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 occupational choice \( O_t \). The changes to the worker’s beliefs in the next period comes through two channels, skill accumulation and learning.

### 7.1 Skill Accumulation

Given the choice of occupation \( \tau_t \) the worker’s true cognitive skills will deterministically increase by

\[
\mathcal{T}_C (O_t) = r_C (1 - q_1 C) + R_C \tau_{Ct} (1 - q_2 C)
\]

and true manual skills by

\[
\mathcal{T}_M (O_t) = r_M (1 - q_1 M) + R_M \tau_{Mt} (1 - q_2 M)
\]

where the \( r \) parameters reflect skill increases that do not depend on tasks, the \( R \) control how quickly skills increase with tasks, and the \( \alpha \) adjust the marginal returns to tasks. The \( q \) terms control the exogenous and endogenous growth rates of human capital over time. Practically, these terms are required for concave life cycle wage profiles. Further generalization of the skill accumulation processes (e.g. a second power term in tasks) did not assist in model fit.

### 7.2 Endogenous Learning

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

\[
Z_{Ct} \equiv C_t^* + \frac{\varepsilon_{Ct}}{\sqrt{e_C \cdot \tau_{Ct}}}, \varepsilon_{Ct} \sim N(0,1)
\]

and

\[
Z_{Mt} \equiv M_t^* + \frac{\varepsilon_{Mt}}{\sqrt{e_M \cdot \tau_{Mt}}}, \varepsilon_{Mt} \sim N(0,1)
\]

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 \( e_C \) and \( e_M \) parameters govern the signal to noise ratio of the \( Z \). 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 \( \varepsilon \). This signal process was originally studied in a two-period, one-dimensional version by Antonovics & Golan (2012). Here the multi-dimensionality is not adding anything because of the independence.
assumptions; it is the single-dimensional model twice. While this is possible to generalize, both signals are already unobservable and any correlation is an additional parameter to estimate.

Since the worker’s current beliefs about his true skills are normal and the noise $\epsilon$ is normally distributed, from the worker’s perspective the $Z_t$ are both normally distributed. It is simple to show that choosing higher-task occupations does not change the expected value of the signal, but it does increase its precision. 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 skills will be normally distributed since the prior beliefs $B_t$ are normal distributions and the $Z$ signals are also normally distributed. For any skill (suppressing skill subscripts), if current beliefs are $B_t \sim N(\mu_t, \frac{1}{p_t})$, the updating rule for average beliefs upon receiving a signal $Z_t$ is a weighted average of the prior average and the signal, and posterior precision is deterministically increasing:

$$
\mu_{(t+1)} = \frac{r \cdot \tau_t Z_t + \mu_t p_t}{r \cdot \tau_t + p_t}
$$

$$
p_{(t+1)} = p_t + r \cdot \tau_t
$$

The weight placed on the signal is a function of the current uncertainty about the skills, the “noise” of the signal as dictated by the parameter $r$, 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.

The model is solved numerically using standard techniques. For details, see Appendix C.

8 Estimation

The model parameters are estimated using Indirect Inference (Gourieroux & Monfort, 1997). I generate simulated occupational choices and wage outcomes from the model and minimize the distance between the simulated and observed outcomes. This approach has both computational and identification advantages. On the computational side, it does not require numerical integration over the unobserved beliefs and shocks that workers are using for their choices and belief updating. For identification, I can ensure that the model replicates key aspects of the data that I discussed in the descriptive statistics and identification argument above, particularly the rate of return moves.

The Indirect Inference approach used here estimates the structural parameters $\theta$ by mini-
mizing a criterion function $Q$:

$$\hat{\theta} = \arg\min_{\theta} Q(\theta) = \arg\min_{\theta} \left( h_S(\theta) - \hat{h} \right)' W \left( h_S(\theta) - \hat{h} \right)$$

(23)

where $\hat{h}$ is the stacked vector of coefficients from regressions 1-11 below using the true data and $h_S(\theta)$ are the coefficients from those same regressions calculated off of simulated data with structural parameters $\theta$. The number of simulations $S$ affects the standard errors of the estimated parameters. $W$ is a weighting matrix which is typically estimated using a two-step procedure as in Gourieroux & Monfort (1997); I use the second step optimal weight matrix.

### 8.1 Matched Coefficients

This section presents the regressions to be matched and their ability to identify the underlying model parameters. The first set of regressions aims to match the relationship between wages and task choices, the second to match cross-sectional worker task choices over time, and the third replicates within-worker transition patterns across tasks. While there is not an exact correspondence, the first set of regressions are most closely related to the structural wage equation, the second to the skill accumulation parameters, and the third to the skill uncertainty parameters. A complete list of the coefficients matched in estimation is given by the first column of Table 8.

The first set of regressions are the conditional averages of wages as a function of years since labor market entry, $t$. Specifically, for each worker-year observation I run an OLS (unless otherwise specified) regression of:

1. Wages onto $t$ and $t^2$.
2. Wages onto $t$, $t^2$, specialization, and magnitude.
3. Fixed effects: Wages onto $t$, $t^2$, specialization, and magnitude.

Matching these regressions ensures that workers in the model have the same average rate of wage growth as the data, as well ensuring the same cross-sectional and within-career relationship between task choices and wages as the data. The difference between regressions (2) and (3) is informative about how much the observed OLS “return” to specialization and magnitude is due to higher wage workers being in those sectors their whole careers. The results have been discussed already in the Descriptive Statistics section.
The second set of regressions are the aggregate patterns in task choices over time. Again, for each worker-year observation I regress

(4) Specialization onto \( t \) and \( t^2 \).

(5) Magnitude onto \( t \) and \( t^2 \).

(6) Squared residuals from Regression 4 onto \( t \) and \( t^2 \).

(7) Squared residuals from Regression 5 onto \( t \) and \( t^2 \).

These regressions ensure that, on average, workers sort across occupations in a way that is consistent with the data. Regressions (4) and (5) look at changes in the average specialization and magnitude measure over time. Regressions (6) and (7) are consistent estimators for the residual variance of the specialization and magnitude measures net of time effects. The general pattern is as discussed above: workers increase their magnitude of tasks over time and move slightly towards more cogitively-specialized jobs.

The final set of regressions focuses on within-worker mobility measures across time:

(8) Absolute value of one-period changes in specialization onto \( t \) and \( t^2 \).

(9) Absolute value of one-period changes in magnitude onto \( t \) and \( t^2 \).

(10) Proportion of the total move size made up of specialization change onto \( t \) and \( t^2 \).

(11) Dummy variable, 1 if “return move”, 0 otherwise, onto \( t \) and \( t^2 \). Conditional on occupational move.

Regressions (8) and (9) summarize the degree of occupational mobility is happening as a function of time. This is the model analogue to the moment “Per-year probability of occupational change.” But now the potential change has two dimensions: workers may make large occupational moves in terms of specialization from their previous job, may make large moves in terms of magnitude, or perhaps both. The results show that, as expected, both types of moves decline over the career.

Regression (10) looks at the proportion of each move that is composed by changes in specialization rather than magnitude. The dependent variable as

\[
p_\theta \equiv \frac{\|\Delta \theta\|}{\|\Delta \theta\| + \|\Delta \tau\|}
\]
where $\Delta$ denotes year-to-year changes. In other words, if workers only ever increase their task magnitude, $p_0$ is always 0, etc. While we know that on average both magnitude and cognitive specialization increase over the career, it could be that within workers these things happen at different times. For instance, in the spirit of Neal (1995) and Pavan (2011) workers could change specialization early in their career to find their best “career” and then increase their magnitude from there. The estimates here are consistent with that story: early in their careers workers make larger specialization changes than magnitude changes, but by 12 years after entry they are making larger task magnitude changes.

Regression (11) has as its dependent variable whether or not the last occupational move was a “return move” as defined in the Identification section. To be specific, a return move is defined as a 1 if $\text{sign} [\Delta \theta_t] \neq \text{sign} [\Delta \theta_{t-1}]$ and a 0 otherwise, where $-1$ indicates the period of the last move. So for example, a series of occupational moves between three occupations such as “20% cognitive tasks $\Rightarrow$ 40% cognitive tasks $\Rightarrow$ 30% cognitive tasks” is a return move, whereas if the last move were to a 60% cognitive job it would not be a return move.

The results from regression (11) may be surprising: most moves are return moves. At labor market entry, around 70% of valid move pairs resulted in a return move. Note one important qualification from the Measurement Error section earlier: I dropped three-digit occupational transitions of the form “Occupation A $\rightarrow$ Occupation B $\rightarrow$ Occupation A” because of concern that these changes were spurious. All of these would be coded return moves, but even after dropping them the majority of moves are return moves. 70% may seem high, but either workers moving randomly across occupations or very noisy measurement of tasks would have up to 50% return moves. Still, the simulation results will show that 70% is far too high for only the skill accumulation model to match.

9 Results

Parameter estimates and confidence intervals are shown in Table 7. Most of the parameters are estimated precisely, but it is difficult to directly interpret many of the magnitudes. Of particular note, the model estimates that workers have no additional information about their initial cognitive skills, but a fairly accurate assessment of their manual skills. Of course over the first few years in the labor market workers gain quite a bit of information about their cognitive skills, so these parameter estimates do not imply workers are still ignorant of their skills at any time after entry. Further interpretation of the parameter estimates is easier by looking at measures
of model fit.

To more clearly demonstrate what the forces can do alone and together to explain the data, I estimate two partial versions of the model and then the full model: a partial version with only skill accumulation, another partial version with only skill uncertainty, and lastly the full model with both forces. The skill accumulation partial model can explain wage growth across the career quite easily, but it fails to explain patterns in worker occupational mobility, in particular return moves. The skill uncertainty partial model can get average wage growth over the career consistent with the data, but cannot explain the full cross-sectional patterns of wages, age, and task choices. The full model is able to fit the regressions significantly better than the separate models. Of course, the fit of the model always improves with an increase in the degrees of freedom, but more importantly the full model is able to simultaneously fit both aggregate wage growth and within-worker occupational mobility, which neither of the individual models could do.

The best-fit simulated regression coefficients from estimating the first partial model, with skill accumulation but no skill uncertainty, are shown in Column 2, Table 8. The skill accumulation partial model can do an excellent job fitting the two OLS wage regressions with different controls in regressions 1 and 2 as well as the fixed effect regression in specification 3. Fitting regression 1 is trivial since the model includes an exogenous wage growth process. Attributing all wage growth to that process, however, would fail to match the regressions that show that task choices are correlated with wages conditional on time since entry. Instead, the model can fit regressions 2 and 3 because of the selection of individuals with different skill levels and mixes into different tasks. On the other hand, the skill accumulation partial model performs poorly for some individual-level worker task transitions. In particular, as argued in the identification section, this model cannot match the rate of return moves. Only 16% of moves early in the career are return moves, compared to 73% in the data.

The best-fit simulated regression coefficients from estimating the second partial model, with skill uncertainty but without skill accumulation, are shown in Column 3, Table 8. Even though workers’ actual skills are fixed, this partial model can still generate wage growth patterns consistent with the data due to experimentation incentives. To fit the wage regressions, however, this partial model does require the exogenous wage process to play a larger role than in the skill accumulation only version. This can be seen since while the model does well predicting wages, it does less well predicting the average task choices associated with that growth. In contrast with the skill accumulation partial model, however, the skill uncertainty partial model can generate
63% of early career moves as return moves, quite close to the 73% from the data. Again, this is consistent with the identification argument above that the number of these moves can help the model identify the relative importance of skill accumulation and skill uncertainty.

The best-fit simulated moments from estimating the full model are shown in Column 4, Table 8. As might be hoped, the full model combines the virtues of both the separate models. The wage regressions are overall fit slightly better, although some individual parameters miss more in order to gain fit on the mobility rates. Like the skill accumulation model, the full model fits the aggregate task choices over time well, and like only the skill uncertainty only model, the full model can match the rate of return moves. Additionally, each of the individual forces under-predicted the year-to-year mobility rates in order to match the rest of the moments, so including both forces fits both those rates better.

10 Counterfactuals

To determine how much of wage growth and task mobility is driven by each force, I simulate the full model twice: first, using the estimated full model parameters but with no skill accumulation, and second using the full model parameters but no skill uncertainty.

The second column of Table 9 shows the counterfactual wage patterns when the skill accumulation parameters are set to 0. Workers then have fixed skills over their careers, but they still are uncertain about their initial levels. The result is that total career earnings go down around 20%, and the linear coefficient on years in the labor force drops by half.

Column 3 of Table 9 does the opposite exercise: workers now know their skills exactly but still gain skills at the baseline rates. The aggregate wage growth numbers are nearly unchanged from the baseline: total earnings over the career go up by about 1%. The intuition for the small estimate of the welfare importance of skill uncertainty requires looking at changes in task choices as well, since tasks and wages are closely linked in the model.

While skill uncertainty does not have a significant net effect on earnings, it plays a significant role in worker occupational mobility. Column 2 of Table 10 shows summary statistics for mobility measures when there is skill uncertainty but no skill accumulation; Column 3 shows summary statistics for when there is only skill accumulation but no skill uncertainty. As seen in Column 2, even with only skill accumulation there is still a large amount of occupational mobility. Even with removing the skill accumulation incentives to switch jobs, the counterfactual model generates around 20% of the baseline rate of per-year $\theta$ mobility and 90% of the per-year
∥τ∥ mobility rate. There is also a change in whether specialization or magnitude moves happen more often: the last row of Table 10 shows that on average moves are composed of more specialization changes in the skill accumulation case than in the skill uncertainty case.

The small net effects of skill uncertainty on wages are the result of the changes in worker task choices. In the model, occupational mobility acts as a hedge: experimenting early in the career can prevent future bad news about one's skills. But the fact that wages barely grow even with full information about skills implies that the hedging is very effective: workers are able to self-insure against adverse new information through occupational choices to the degree that they can eliminate almost all potential wage losses from skill uncertainty. Of course this is not necessarily true as a matter of theory, but is true of the model at the estimated parameters. The same is not true with skill accumulation: the model shows that in the absence of skill growth, workers will take higher wage jobs earlier in their career but also still face large career earnings losses.

The amount of hedging workers do through occupational choice can be quantified by running additional counterfactuals. First, I simulate a partial model with only skill accumulation and no skill uncertainty. Then, I simulate the full model, but force workers to choose the same occupations over time they did in the partial model. Comparison of these two simulated models allows us to determine the wage effects of uncertainty while removing the ability of the workers to react to that uncertainty through changing their occupational choices. Theory implies that workers must be weakly better off if they can freely adjust their occupations, but it does not say quantitatively how well they can use occupational choices to self-insure.

The results from these counterfactuals are shown in Table 11. Recall that in the earlier counterfactuals if workers lose perfect certainty it reduces earnings over the career by around 1% of lifetime earnings, around $20,000. On the other hand, if workers lose perfect certainty but also cannot react to that loss of certainty through choosing occupations better suited to help them learn their skills, their wages over their career go down by 2.5% – a total welfare loss of around $50,000. So when workers facing skill uncertainty are able to shift their occupational choices in response to that uncertainty in order to learn more quickly, they regain around $30,000 dollars of that lost welfare. In the end, mobility and earnings are closely related: It is not simply that that if skill uncertainty were somehow eliminated, workers would earn just a bit more and move occupations less often. Instead, the reason workers are earning so much now even with uncertainty is that they are able to change occupations often to protect themselves from the uncertainty.
The worker mobility counterfactuals from Table 10 give some indications that the two forces do not act independently in terms of mobility. In the first row, eliminating either learning or human capital gets rid of around 75% of observed mobility across specializations. In other words, for this particular type of mobility, the two forces have a positive interaction: their combination creates more mobility across specializations than the sum of the two separately. On the other hand, the second row shows that mobility across magnitudes can be divided independently: learning accounts for \( \frac{6}{7} \) of mobility across magnitudes and human capital accounts for \( \frac{1}{7} \). With such a low estimated rate of manual human capital accumulation, when an individual moves upwards in both cognitive and manual tasks it must be the case that they learn about their initial ability instead of gaining more skills.

Finally, the combination of human capital and learning actually works against return moves. As the third row of Table 10 shows, the level of return moves with just learning is higher than that with both forces. Intuitively, there are situations where a worker learns they have some skills and that learning would induce a return move, but the amount of human capital they accumulated between periods means it is no longer worth it to make that move. This is similar to the theory of firm-specific human capital, where human capital can inhibit moves due to learning about the firm-worker match, as in Nagypal (2007): the secretary may learn he was a bad fit for the firm all along, but now that he knows the filing system perfectly there is no point in moving.

In summary, the counterfactuals show that wage growth by itself can be explained in terms of considering workers increasing their skill levels and moving up a career ladder. On the other hand, a significant part of early career mobility – even mobility that involves large changes in what workers are actually doing – can be explained as a wage-maximizing response to worker uncertainty.

11 Conclusion

In this paper, I constructed and estimated a model of worker choice that includes both skill accumulation and skill uncertainty. The end goal was to distinguish the forces of skill accumulation from skill uncertainty on wage growth and occupational choices. Typically these two economic forces had been looked at in isolation, but their relative importance had not been quantified. My results indicate that skill accumulation is much more significant than skill uncertainty for wage growth. However, I also found that the estimated small effects of skill un-
certainty are due to workers responding to the uncertainty through their occupational choices. This suggests that in labor markets with less flexibility than the cohort studied here, workers entering the US labor market around 1984, the welfare costs of uncertainty could be more significant.

There are a number of further steps to take towards a more realistic picture of the skill accumulation and skill uncertainty processes. First, there are remaining measurement issues that can be improved through use of new data. For example, worker-level measurement of tasks can dramatically reduce the measurement error in tasks and allow for a more precise characterization of task transitions, especially within occupations. Worker-level task data could also help find the best parametrization of the wage equation and learning process, which at this point require a number of assumptions given the noisiness of the data.

Second, there are extensions to the model that could change the welfare analysis significantly. For example, the model assumes wage maximizing workers. Many of the things we consider “learning” in the labor market are learning about one’s own preferences, such as the non-pecuniary match with a particular occupation or firm, rather than simply about wages. Additionally, there are no labor market frictions in the model, which given the results may be important in evaluating the welfare consequences of learning. While these modeling extensions would be difficult to add to an already computationally and data intensive model, perhaps descriptive evidence can be found that some of these features are as important as the ones modeled here.

**Bibliography**


Computational Bibliography


A Figures and Tables

Figure 1: Example O*NET Question

5. Mathematics

Using mathematics to solve problems.

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

<table>
<thead>
<tr>
<th>Not Important*</th>
<th>Somewhat Important</th>
<th>Important</th>
<th>Very Important</th>
<th>Extremely Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

* 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?

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

1  2  3  4  5  6  7

Highest Level
Figure 2: Task Scores, White-Collar and Blue-Collar Occupations
Figure 3: Task Scores of Three-Digit Occupations by One-Digit Occupational Classification
Figure 4: Specialization and Magnitude
Figure 5: Identification, occupational choices

(a) Sample Career, Human Capital Accumulation Only

(b) Sample Careers, Learning Only
Table 1: Cross-Sectional Worker Characteristics Over Time

<table>
<thead>
<tr>
<th>Year Since Entry</th>
<th>Mean Wages</th>
<th>Std. Dev. Wages</th>
<th>Mean $\theta$</th>
<th>Mean $|\tau|$</th>
<th>Std. Dev. $\theta$</th>
<th>Std. Dev. $|\tau|$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12.0</td>
<td>5.4</td>
<td>.23</td>
<td>-.30</td>
<td>1.03</td>
<td>.84</td>
</tr>
<tr>
<td>5</td>
<td>13.9</td>
<td>5.9</td>
<td>.12</td>
<td>-.15</td>
<td>0.95</td>
<td>.83</td>
</tr>
<tr>
<td>10</td>
<td>16.6</td>
<td>8.2</td>
<td>.04</td>
<td>-.09</td>
<td>0.97</td>
<td>.88</td>
</tr>
<tr>
<td>15</td>
<td>17.1</td>
<td>9.0</td>
<td>.03</td>
<td>-.02</td>
<td>1.03</td>
<td>.95</td>
</tr>
<tr>
<td>20</td>
<td>18.8</td>
<td>10.9</td>
<td>-.16</td>
<td>.22</td>
<td>.96</td>
<td>1.18</td>
</tr>
</tbody>
</table>

Sample size of 1774 white males with exactly 12 years of schooling. On average, 12.9 valid yearly observations per person.

Table 2: OLS Regression, Monthly Level, Dependent Variable: Hourly Real Wage

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>16.6</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$|\tau|$</td>
<td>1.41</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-0.20</td>
<td>(0.06)</td>
</tr>
</tbody>
</table>

$N = 21,647$. $R^2 = 0.03$. Both $\theta$ and $\|\tau\|$ have mean 0, std. dev. 1.

Table 3: Fixed Effects Regression, Monthly Level, Dependent Variable: Hourly Real Wage

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>16.6</td>
<td>(0.15)</td>
</tr>
<tr>
<td>$|\tau|$</td>
<td>.95</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-0.27</td>
<td>(0.06)</td>
</tr>
</tbody>
</table>

$N_i = 1,766$. $R^2 = .03$. Average obs per group = 12.3
Both $\theta$ and $\|\tau\|$ have mean 0, std. dev. 1.
Corr($u_i, X\beta$) = .07
Table 4: Occupational Moves

4(a): Predicted Monthly Probability of Change of O*NET Occupation Name, Linear Probability Model

<table>
<thead>
<tr>
<th>Time Since Entry, Months</th>
<th>Pr(Occupation Change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.3%</td>
</tr>
<tr>
<td>60</td>
<td>4.2%</td>
</tr>
<tr>
<td>120</td>
<td>3.3%</td>
</tr>
<tr>
<td>240</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

4(b): Predicted Monthly Average Changes in $\theta$ and $\|\tau\|$

<table>
<thead>
<tr>
<th>Time Since Entry, Months</th>
<th>Avg. Change in $\theta$</th>
<th>Avg. Change in $|\tau|$</th>
<th>Avg. Size of Change in $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.26%</td>
<td>0.26%</td>
<td>4.8%</td>
</tr>
<tr>
<td>60</td>
<td>-0.21%</td>
<td>0.22%</td>
<td>4.1%</td>
</tr>
<tr>
<td>120</td>
<td>-0.17%</td>
<td>0.17%</td>
<td>3.4%</td>
</tr>
<tr>
<td>240</td>
<td>-0.07%</td>
<td>0.08%</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

All changes are stated in terms of % of one standard deviation.

Table 5: Return Moves

OLS Regression, Dependent Variable: “Return Move” Dummy Variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff. Est.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>(.01)</td>
</tr>
<tr>
<td>Time Since Entry (Months)</td>
<td>-3.01 x 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>(.0005)</td>
</tr>
</tbody>
</table>

Table 6: One-Digit Occupation Averages, 10 Years After Entry

<table>
<thead>
<tr>
<th>Occupational Category</th>
<th>% of Workers</th>
<th>Wage</th>
<th>$\theta$</th>
<th>$|\tau|$</th>
<th>Firm Tenure (Months)</th>
<th>AFQT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional and Technical</td>
<td>6</td>
<td>16.2</td>
<td>-0.66</td>
<td>0.88</td>
<td>58</td>
<td>48</td>
</tr>
<tr>
<td>Managers and Administrators</td>
<td>6</td>
<td>15.9</td>
<td>-1.09</td>
<td>0.63</td>
<td>50</td>
<td>48</td>
</tr>
<tr>
<td>Sales</td>
<td>1</td>
<td>18.4</td>
<td>-1.41</td>
<td>0.91</td>
<td>38</td>
<td>57</td>
</tr>
<tr>
<td>Clerical</td>
<td>22</td>
<td>15.2</td>
<td>-0.02</td>
<td>-0.69</td>
<td>53</td>
<td>48</td>
</tr>
<tr>
<td>Clerical</td>
<td>23</td>
<td>19.4</td>
<td>-1.14</td>
<td>0.22</td>
<td>52</td>
<td>43</td>
</tr>
<tr>
<td>Operatives</td>
<td>13</td>
<td>16.1</td>
<td>1.01</td>
<td>0.06</td>
<td>54</td>
<td>38</td>
</tr>
<tr>
<td>Transport Equipment Operatives</td>
<td>5</td>
<td>15.7</td>
<td>0.08</td>
<td>-0.13</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>Laborers</td>
<td>10</td>
<td>16.3</td>
<td>0.33</td>
<td>-0.32</td>
<td>51</td>
<td>34</td>
</tr>
<tr>
<td>Farmers and Farm Laborers</td>
<td>2</td>
<td>11.2</td>
<td>-0.26</td>
<td>-0.22</td>
<td>75</td>
<td>38</td>
</tr>
<tr>
<td>Service Workers</td>
<td>12</td>
<td>12.5</td>
<td>-0.40</td>
<td>-0.12</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Private Household Workers</td>
<td>0.3</td>
<td>9.3</td>
<td>-0.24</td>
<td>-0.35</td>
<td>51</td>
<td>31</td>
</tr>
</tbody>
</table>

$\theta$, $\|\tau\|$ have mean 0 and std. dev. 1. AFQT has mean 45 and std. dev. 27.

Table 7: Parameter Estimates

<table>
<thead>
<tr>
<th>Cognitive</th>
<th>Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0, a_1$</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>(0.041,0.074)</td>
</tr>
<tr>
<td>$k$</td>
<td>6.898</td>
</tr>
<tr>
<td></td>
<td>(6.547,7.248)</td>
</tr>
<tr>
<td>$R$</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>(0.116,0.137)</td>
</tr>
<tr>
<td>$q_1, q_2$</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>(0.632,1.144)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.374</td>
</tr>
<tr>
<td></td>
<td>(0.335,0.414)</td>
</tr>
<tr>
<td>$r$</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>(−0.367,0.518)</td>
</tr>
<tr>
<td>$e$</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.009,0.015)</td>
</tr>
<tr>
<td>$\sigma_0^2$</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.909,1.090)</td>
</tr>
<tr>
<td>$w_g$</td>
<td>−6.953</td>
</tr>
<tr>
<td></td>
<td>(−7.229,−6.677)</td>
</tr>
<tr>
<td>$w_{gr}$</td>
<td>0.370</td>
</tr>
<tr>
<td></td>
<td>(0.364,0.375)</td>
</tr>
</tbody>
</table>

Asymptotic 95% Confidence Intervals in parentheses. The * indicates that at the estimated value of $r_m$, $q_{1m}$ is unidentified.
<table>
<thead>
<tr>
<th></th>
<th>Data Estimate</th>
<th>Only HC</th>
<th>Only Learning</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLS: Average Wages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.6</td>
<td>12.7</td>
<td>12.8</td>
<td>12.7</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>0.36</td>
<td>0.35</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>$-2.7 \times 10^{-3}$</td>
<td>$-4.7 \times 10^{-3}$</td>
<td>0.03</td>
<td>$-6.4 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>OLS: Average Wages with Tasks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.9</td>
<td>12.9</td>
<td>12.8</td>
<td>12.9</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>0.35</td>
<td>0.34</td>
<td>0.39</td>
<td>0.36</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>$-3.2 \times 10^{-3}$</td>
<td>$-4.7 \times 10^{-3}$</td>
<td>0.02</td>
<td>$-6.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>$-9.3 \times 10^{-3}$</td>
<td>0.30</td>
<td>-0.02</td>
<td>0.23</td>
</tr>
<tr>
<td>$|\tau|$</td>
<td>1.12</td>
<td>1.22</td>
<td>1.14</td>
<td>1.08</td>
</tr>
<tr>
<td><strong>Wage Regression w/ Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.1</td>
<td>12.0</td>
<td>12.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>0.41</td>
<td>0.36</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>$-3.3 \times 10^{-3}$</td>
<td>$-5.3 \times 10^{-3}$</td>
<td>0.02</td>
<td>$-6.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.91</td>
<td>0.68</td>
<td>0.88</td>
<td>0.55</td>
</tr>
<tr>
<td>$|\tau|$</td>
<td>0.70</td>
<td>0.96</td>
<td>1.03</td>
<td>1.30</td>
</tr>
<tr>
<td>$\theta \times$ Years Since Entry</td>
<td>-0.07</td>
<td>$-5.4 \times 10^{-3}$</td>
<td>-0.02</td>
<td>$-3.6 \times 10^{-3}$</td>
</tr>
<tr>
<td>$|\tau| \times$ Years Since Entry</td>
<td>-0.02</td>
<td>0.01</td>
<td>$5.4 \times 10^{-3}$</td>
<td>$-5.9 \times 10^{-4}$</td>
</tr>
<tr>
<td><strong>OLS: Average $\theta$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.26</td>
<td>0.44</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>-0.02</td>
<td>-0.06</td>
<td>$8.9 \times 10^{-3}$</td>
<td>-0.04</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>$6.6 \times 10^{-5}$</td>
<td>$1.3 \times 10^{-3}$</td>
<td>$1.5 \times 10^{-4}$</td>
<td>$7.5 \times 10^{-4}$</td>
</tr>
<tr>
<td><strong>OLS: $\sigma^2_\theta$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.97</td>
<td>1.19</td>
<td>0.61</td>
<td>1.13</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>$6.9 \times 10^{-4}$</td>
<td>-0.01</td>
<td>0.02</td>
<td>$-6.8 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>OLS: Average $|\tau|$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.25</td>
<td>-0.30</td>
<td>0.02</td>
<td>-0.28</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>-0.01</td>
<td>0.02</td>
<td>$-1.8 \times 10^{-3}$</td>
<td>0.02</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>$4.7 \times 10^{-4}$</td>
<td>$-3.5 \times 10^{-4}$</td>
<td>$3.4 \times 10^{-5}$</td>
<td>$-3.4 \times 10^{-4}$</td>
</tr>
<tr>
<td><strong>OLS: $\sigma^2_|\tau|$$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.47</td>
<td>1.13</td>
<td>0.60</td>
<td>0.66</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>0.04</td>
<td>$-3.5 \times 10^{-3}$</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>OLS: Size of Changes in $\theta$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.49</td>
<td>0.31</td>
<td>0.23</td>
<td>0.43</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>$-1.0 \times 10^{-3}$</td>
<td>-0.01</td>
<td>$7.0 \times 10^{-3}$</td>
<td>-0.01</td>
</tr>
<tr>
<td><strong>OLS: Size of Changes in $|\tau|$$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.37</td>
<td>0.20</td>
<td>0.25</td>
<td>0.28</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>$-2.6 \times 10^{-4}$</td>
<td>$-7.0 \times 10^{-3}$</td>
<td>$-7.7 \times 10^{-3}$</td>
<td>$-8.4 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>LPM: Return Moves</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.73</td>
<td>0.16</td>
<td>0.63</td>
<td>0.53</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>$-3.6 \times 10^{-3}$</td>
<td>$-4.4 \times 10^{-3}$</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td><strong>OLS: Proportion of Changes Due to $\Delta \theta$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.57</td>
<td>0.76</td>
<td>0.54</td>
<td>0.56</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>$-5.9 \times 10^{-3}$</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

Criterion Value

276 221 72
### Table 9: Wage Growth Decomposition

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Only Learning</th>
<th>Only HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Wage</td>
<td>12.8</td>
<td>12.8</td>
<td>13.1</td>
</tr>
<tr>
<td>Years Since Entry</td>
<td>0.38</td>
<td>0.18</td>
<td>0.39</td>
</tr>
<tr>
<td>Years Since Entry$^2$</td>
<td>-0.006</td>
<td>-0.003</td>
<td>-0.007</td>
</tr>
<tr>
<td>Average Hourly Wage, Years 0-40</td>
<td>17.1</td>
<td>13.6</td>
<td>17.3</td>
</tr>
<tr>
<td>% of total wages due to growth</td>
<td>25%</td>
<td>15%</td>
<td>24.5%</td>
</tr>
</tbody>
</table>

### Table 10: Counterfactual Task Mobility

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Only Learning</th>
<th>Only HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-Year $\theta$ Mobility</td>
<td>0.43</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>Per-Year $|\tau|$ Mobility</td>
<td>0.28</td>
<td>0.24</td>
<td>0.04</td>
</tr>
<tr>
<td>Return Moves</td>
<td>0.53</td>
<td>0.61</td>
<td>0.20</td>
</tr>
<tr>
<td>% of Moves Made up of $\theta$ Switches</td>
<td>0.56</td>
<td>0.31</td>
<td>0.75</td>
</tr>
</tbody>
</table>

### Table 11: Counterfactual Welfare, Hedging vs. No Hedging

<table>
<thead>
<tr>
<th></th>
<th>% of Lifetime Welfare Loss</th>
<th>$\text{Value of Welfare Loss}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference Case: Perfect Certainty</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Add Uncertainty, Force Reference Occupational choices</td>
<td>2.5%</td>
<td>49,500</td>
</tr>
<tr>
<td>Add Uncertainty, Allow New Occupational Choices</td>
<td>1%</td>
<td>19,000</td>
</tr>
</tbody>
</table>

## B Sample Selection and Occupational Merging

### NLSY Sample Selection

I restrict the NLSY sample to white males from the core sample with a high school degree and no college attendance 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 between all employers. Workers who work full time at least 20 of the past 24 weeks are considered attached to the labor force. The time of labor market entry is set as the first week where the worker is attached to the labor force after leaving school, so no earlier than 6 months after leaving school. Additionally, I drop workers who are ever in military service.

### Forming the O*NET Task Measures

My procedure for creating the cognitive and manual tasks scores is slightly different than what has been used in the literature, e.g. Poletaev & Robinson (2008) and Yamaguchi (2012). Both
their procedures and mine essentially create cognitive and manual scores as a linear combination of the responses to a number of questions taken from the O*NET data. However, they use Principal Component Analysis to recover the weights on the different questions, where I use weights determined by their coefficients in a wage regression from the NLSY. The end results are similar for occupation rankings in terms of cognitive and manual tasks; comparison of different methods for forming these scores remains an area for future research.

The following questions were taken from the “Work Activities” section of the O*NET survey. Workers were asked to rate the “level” and “importance” of each of these activities in their job. The two scores are correlated over 0.95, so only the level score was used.

- **Cognitive:**
  - Getting Information
  - Processing Information
  - Analyzing Data or Information
  - Making Decisions and Solving Problems
  - Thinking Creatively
  - Updating and Using Relevant Knowledge
  - Developing Objectives and Strategies
  - Organizing, Planning, and Prioritizing Work

- **Manual:**
  - Performing General Physical Activities
  - Handling and Moving Objects
  - Controlling Machines and Processes
  - Operating Vehicles, Mechanized Devices, or Equipment
  - Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment
  - Repairing and Maintaining Mechanical Equipment Repairing and Maintaining Electronic Equipment
I then linearly projected wage onto worker characteristics and survey responses for their chosen occupation in each year using the regression

\[
E [ \text{Real Wage}_{it} | \text{Covariates}_{it} ] = \beta_{0i} + \beta_1 \text{Urban/Rural Dummy}_{it} + \\
\beta_2 t + \beta_3 t^2 + \beta_4 \text{Actual Labor Market Experience}_{it} + \\
\beta_5 \text{Actual Labor Market Experience}^2_{it} + \beta_6 \text{Number of Children}_{it} + \beta_7 \text{Marital Status Dummy}_{it} + \\
\beta_C \cdot \text{Cognitive Task Responses}_{it} + \beta_M \cdot \text{Manual Task Responses}_{it}
\]

The regression is estimated using fixed effects, which allows for individual-specific intercept terms (the \(\beta_{0i}\)). I then form the scores for each occupation using the estimated coefficients from the wage equation:

\[
\hat{\tau}_C j = \hat{\beta}_C \cdot \text{Cognitive Task Responses in Occupation } j
\]

\[
\hat{\tau}_M j = \hat{\beta}_M \cdot \text{Manual Task Responses in Occupation } j
\]

and normalize each score between 0 and 1 to get the final \(\tau_C\) and \(\tau_M\) for each occupation. All these regressions and scores are available upon request.

**Merging the O*NET with the NLSY**

Between the two surveys there are three different occupational coding schemes: the Census 1970 scheme in the NLSY from 1979 to 2000, the Census 2000 scheme from 2002 to 2008, and the SOC 2000 categories in O*NET.

In the NLSY, from the initial survey until 2000 occupations were classified according to the US Census 3-digit 1970 classifications\(^8\). From 2002 onwards the classification was updated to use the US Census 3-digit 2000 classification, which restructured, renamed, and recoded almost every occupation. The O*NET uses SOC (Standard Occupational Classification) 2000 codes for all of its occupational level data. There is not a 1-to-1 correspondence between any of these coding systems.

To merge in the occupations then required crosswalks that could take a Census 1970 code or a Census 2000 code a return a SOC code. For the Census 1970 classifications to SOC categories I created a crosswalk by hand by comparing the titles and content of the occupations. In a

---

\(^8\)From 1982-2000 the NLSY also records occupation under the 1980 coding system, but I do not use it since it misses the beginning of the sample.
number of cases there was not an obvious unique match; for these I coded all the possible matches and then chose the “most plausible” one by taking the one with highest employment share.

The Census 2000 classification to SOC crosswalk was based on a crosswalk provided by the National Crosswalk Service but required a significant amount of manual adjustment before it matched the SOC codes.

Both crosswalks are available on request.

C Computation

To solve the model, I used backwards interpolation of the value functions with numerical integration and interpolation routines from the NumPy (see Van Der Walt et al. (2011)) and SciPy (see Jones et al. (2001)) packages of the Python programming language. In addition, for parameter estimation I used the NLOpt package (Johnson, 2012) implementation of the BOBYQA algorithm for bound constrained minimization (Powell, 2009). The pictures and graphs were generated using Matplotlib (Hunter, 2007). See the “Computational Cites” section above for references. All code is available upon request.