A Dynamic Analysis of Educational, Occupational, and Inter-firm Mobility Decisions

Paul Sullivan*

Population Studies Center

University of Michigan

October 2004

Abstract

This research examines educational attainment and mobility between firms and occupations using a dynamic structural model of career choices. The model expands on previous work by jointly modeling transitions between firms and occupations within a model of career choice. Previous work has generally focused on mobility between either firms or occupations in isolation, which ignores the fact that people make occupational choices and job search decisions jointly. Incorporating mobility between firms and occupations within a unified model provides parameter estimates that indicate the relative importance of firm and occupation-specific factors in determining career choices. The estimates suggest that employment choices are driven jointly by firm-specific factors such as matching in wages and occupation-specific factors such as heterogeneity in skills and preferences for different types of work. The estimates also indicate that both firm and occupation-specific human capital play a role in determining wages. The parameters of the dynamic structural model are estimated with maximum simulated likelihood using data on individuals' educational and employment choices from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY).

*I would like to thank my advisors, Steven Stern, John Pepper, and William Johnson for their extremely helpful comments and suggestions. Funding from the Bankard Fund for Political Economy at the University of Virginia and the University of Virginia College of Arts & Sciences is gratefully acknowledged. Email: paulsull@isr.umich.edu.

1 Introduction

Over the course of their careers, people choose how much education to obtain, which occupations to work in, and when to move between occupations and firms. Educational choices are made conditional on future employment plans because educational attainment determines, in part, which occupations are feasible career paths. Many factors enter into employment decisions, but two central considerations are at which firm and in which occupation to work. People frequently move between occupations and firms, and this mobility plays an important role in shaping the path of earnings over time.

This research expands on previous work by jointly modeling transitions between firms and occupations within a dynamic structural model of career choices. Previous work has generally focused on mobility between either firms or occupations in isolation, which ignores the fact that people make occupational choices and job search decisions jointly. Incorporating mobility between firms and occupations within a unified model provides parameter estimates that indicate the relative importance of firm and occupation-specific factors in determining career choices.

The goal of this research is to further the understanding of how people make decisions about educational attainment and job mobility using a model that combines features of human capital and job search models. Individuals in the model maximize their discounted expected utility by choosing when to attend school, when to move between firms and occupations, and when to be unemployed. Future wages depend upon previous choices because workers accumulate education and firm and occupation-specific human capital. One reason people choose to work in different occupations is that they have different abilities and preferences for working in each occupation. Employment decisions also are driven by firm-specific matching in wages and non-pecuniary utility flows. The wage and non-wage match values allow a worker's productivity and valuation of non-wage job characteristics to vary across firms, so workers search for suitable matches with firms across occupations. In addition, workers search for jobs across occupations within their current firm. This type of internal occupational mobility appears to be quite common, especially for younger workers. Transitions between firms and occupations are produced by the interaction of firm-specific match values, occupation-specific skill heterogeneity, human capital, and randomness in job offers and utility shocks.

The importance of understanding the determinants of mobility between firms and occupations is suggested by the frequency of transitions between occupations in the 1979 cohort of the National Longitudinal Survey of Youth (NLSY).[1] In this work occupations are aggregated into five broadly defined groups. In these data, I find that nearly half of the time that a worker moves to a new firm, he also moves to a new occupation. Mobility between occupations within firms is also common. Conditional on being employed at the same firm in consecutive years, there is nearly a 20% chance that a worker will move between occupations within that firm. Mobility between occupations declines steadily with age, which is broadly consistent with either human capital effects on utility flows or the effects of occupation or firm-specific job matching over time. The structural parameter estimates provide information about the relative importance of each of these factors in determining mobility.

The parameters of the structural model are estimated by maximum simulated likelihood using data on the career choices of young men from the NLSY. The likelihood function follows directly from the recursive numerical solution to each individuals' dynamic programming problem and assumptions about the distribution of unobserved heterogeneity. Estimation is computationally expensive for two primary reasons. First, the rich specification of unobserved heterogeneity requires the evaluation of high dimensional integrals when evaluating the expected value functions and the likelihood function. These integrals are approximated using simulation methods. Second, the state space of the model is large, partly because the choice set is large. Agents in the model have a large number of choices in each period because jobs are differentiated by firm and occupation. Also, the inclusion of dual activities such as attending school while working results in a large number of choices available to people in each period. The state space is made even larger by the inclusion of wage and non-wage match values in the state space, in addition to firm and occupation-specific human capital. These complications are addressed by using interpolation methods to reduce the number of times the value functions need to be calculated, and by modeling human capital in a novel way that reduces the size of the state space.

2 Literature Review

The model presented in this paper builds on previous work in the areas of human capital investment, occupational choice, and dynamic labor supply. This section briefly surveys the extensive literature

[1] The levels of within and across firm occupational mobility in the data are discussed in detail in Section 3.

in these areas. Early papers in the human capital literature include Roy (1951), who introduces the concept of self selection in employment choices based on skills. The basic framework of the Roy model is extended by Heckman and Sedlacek (1985) to incorporate self selection in industrial sector choice and by Willis (1996) to include self selection in occupational choices. More recently, Gould (2002) uses this type of framework to examine the relationship between occupational choices and wage inequality. Keane and Wolpin (1997) examine educational and occupational choices within a dynamic framework that allows work experience and education to be accumulated endogenously.

Occupational choices have also been examined within the framework of search and matching models. Miller (1984) estimates a dynamic model where workers learn about how well suited they are to various occupations by working at firms within each occupation.[2] McCall (1990) develops a model of job search where job matching occurs between workers and firms and occupations. In this model it is not possible for workers to switch occupations within a firm. His empirical results suggest that occupational matching exists. Neal (1999) develops a job search model where workers search for suitable matches with both types of work and individual firms, where a type of work is defined using occupation and industry codes. Neal's theoretical model predicts that workers should search for jobs in two stages, first finding a type of work, and then finding a suitable firm within that line of work. He finds that the NLSY data supports the optimality of the two-stage search strategy.

The issue of how people make career choices is closely tied to questions about how human capital affects wages. The fact that wages rise with tenure has been well documented. However, there is still debate about whether this is due to the effect of general work experience, employer-specific experience, industry-specific experience, or occupation-specific experience. Neal (1995) and Parent (2000) provide evidence that human capital is primarily industry-specific. Standard models typically do not include industry-specific tenure, instead they normally include some combination of general and firm-specific experience. More recently, Kambourov and Manovskii (2002) report that after tenure in an occupation has been accounted for, tenure in an industry has little importance in determining wages. This result suggests that career choices should be defined using occupations.

[2] The special structure of the problem allows the parameters to be estimated using the Gittins index solution to the multi-armed bandit problem.The Gittins index cannot be used to solve the model presented in this paper due to the presence of switching costs. See Banks and Sundaram (1994) for details about the Gittins index when switching costs are present. The Gittins index is also not applicable to problems where the assumption of indepence across arms is violated. This is the case in the present model because firm-specific work exerience affects wages in multiple occupations within a firm.

In addition to wage growth due to the accumulation of human capital, wages also increase over the course of a career as a result of transitions between firms. Mincer and Jovanovic (1981) were the first to estimate the magnitude of the increases in wages resulting from job search. Topel and Ward (1992) extend this type of analysis, and conclude that job changes account for at least a third of early career wage growth. Within-firm job changes have received less attention in the empirical literature, and typically have been studied with a focus on promotions rather than the internal occupational mobility examined in this work. For example, McCue (1996) finds that promotions account for about 15% of wage growth for men over the life cycle.

Educational choices have been examined frequently using the human capital framework. Topics of study include estimating the return to investment in schooling and examining the school attendance decision. Willis and Rosen (1979) develop and estimate a model which accounts for self selection in schooling choices based on ability. More recently, schooling choices have been examined with a focus on high school completion by Eckstein and Wolpin (1999), and with a focus on the relationship between borrowing constraints and college attendance by Keane and Wolpin (2001).

Dynamic labor supply models are well suited to formulation as discrete choice dynamic programming problems because this approach allows for the type of forward-looking behavior that plays an important role in employment choices.[3] Labor supply questions investigated using dynamic programming models range from teacher labor supply (Stinebrickner (2001)), to retirement (Rust and Phelan (1997) and Berkovec and Stern (1991)). Berkovec and Stern (1991) show that firm-specific job matching is an important aspect of employment decisions. Mobility between firms is also investigated by Wolpin (1992), who estimates a structural model of labor mobility that focuses on transitions between firms and general and firm-specific human capital accumulation.

The model developed in this work is most closely related to Keane and Wolpin (1997), so I will discuss their work in more detail. Keane and Wolpin estimate a discrete choice dynamic programming model of career decisions using the NLSY. People in the model choose between attending school, remaining at home, working in a white collar job, working in a blue collar job, and working in the military. Firms are not explicitly included in the model, so the model does not capture the effect of firm-specific matching or work experience on wages. The empirical results suggest that heterogeneity and comparative advantages in skill endowments are an important determinant of

[3] See Keane and Wolpin (1994) and Rust (1994) for surveys of the solution and estimation of dynamic programming models.

occupational choices and educational attainment. The model is used to predict the impact of a college tuition subsidy on school completion and occupational choices.

I extend their model along several dimensions. One simple extension is that the number of occupational classifications is increased from two to five. Disaggregating occupations causes occupational mobility to increase substantially, although the level of persistence in occupational choices is still high. The model presented in this paper incorporates the role of individual firms in employment choices by including firm-specific human capital, and firm-specific wage and non-wage matching. The distinction between firms and occupations allows the model to explore the relationship between mobility between firms and occupations and labor market outcomes. Ignoring firm-specific matching and firm-specific human capital may produce an upward bias in estimates of the effect of occupation-specific work experience and skill endowments on wages. The model also allows for dual activities, such as employment while attending school, and incorporates earning a GED as a choice variable.[4]

3 Data

The parameters of the model are estimated using the National Longitudinal Survey of Youth 1979 cohort (NLSY). This data set includes detailed information about the educational and employment experiences of a nationally representative sample of 12,686 men and women who were 14-21 years old when first interviewed in 1979. Data are currently available through 1998, with a retention rate of 84.3%. Interviews were conducted annually up until 1994, and then biennially in the following years. The data provide a rich set of educational information about each respondent, including dates of school attendance and dates of graduation and GED receipt. Employment data include the duration of every employment spell over the sample period, along with the corresponding wages, hours, and occupation for each employment spell. This information allows for the identification of transitions between employers and occupations, as well as the patterns of wage changes over the career.

The NLSY consists of a nationally representative core sample, a military sample, and a supplemental sample that over-samples blacks, Hispanics, and economically disadvantaged whites. This analysis uses only white men from the nationally representative core sample. Individuals who are

[4] See Keane and Wolpin (2001) and Eckstein and Wolpin (1999) for examples of models that also allow for employment while attending school.

older than age sixteen in the first year of the NLSY are not used. Individuals remain in the data set up to age thirty or until the observation is truncated at the first instance of missing information about yearly labor force status or the occupation of a yearly job. Respondents are dropped from the sample if they provide insufficient information to construct a history of educational attainment. Respondents are also dropped from the sample if they ever serve in the military or work as a farmer. The final sample consists of 1,023 men who remain in the sample for an average of 10.37 years, resulting in 10,609 "person years" of data.

The decision period in the model corresponds to a school year, which runs from September to August. The data are aggregated using an approach similar to that of Keane and Wolpin (1997). Yearly school attendance is assigned using detailed information on monthly school attendance and grade completion. The methodology used to assign yearly school attendance consists of several steps. First, the amount of education accumulated by each sample member over the sample period is determined using the variable that indicates the highest grade completed as of each interview year. Then, starting in the first year, individuals are considered to be attending school if they report attending school during the year and completing a grade by the next year. If this approach fails to assign all the accumulated years of education, then the process is repeated using the weaker requirement that the person reports completing a grade or attending school during a year. Receipt of a GED is coded using yearly information on whether or not a person ever earned a GED.

Yearly employment status is determined using the weekly labor force record. The yearly employment activity is the activity (a specific employer or unemployment) in which the most weeks were spent during the year. The number of weeks spent unemployed and employed full time at each employer are counted for each decision year. Jobs consisting of less than twenty hours of work per week are counted as time spent unemployed. The work activity in which the most weeks were spent during the school year is coded as the yearly labor force activity. For example, suppose that during a year a person works at firm A for 22 weeks, works at firm B for 10 weeks, and spends 20 weeks unemployed. The primary activity for this year is working at firm A, so working at firm A is coded as the yearly activity. The yearly occupation is the one corresponding to firm A. Given the assumption that employment is full-time, an individual's wage is converted into a yearly wage by multiplying the hourly wage by 2,000 hours.

Transitions between firms are identified using the NLSY survey variables that indicate whether or not a current employer is the same as an employer in the previous year. One unavoidable consequence of the aggregation of weekly data into yearly data is that yearly data understate the number of transitions between firms. The identification of transitions between firms is a key feature of the model presented in this paper, so it is important to consider the effects of aggregation on the number of transitions between firms present in the data. One way of assessing the effects of aggregation is to compare the average number of jobs that a person holds over the twenty year sample using different levels of aggregation. Using the weekly NLSY employment record, the average number of jobs is 11. When the data are aggregated to half-yearly, the average number of jobs falls to 7. Using yearly data, the average number of jobs is 6. The effects of aggregation are fairly large when moving from weekly to half-yearly data, but relatively small when moving from half-yearly to yearly data.[5]

The NLSY data provides information on occupational codes at the three digit level. The level of detail provided in these codes raises questions about the proper definition of an occupation. The human capital model presented in this paper suggests that an occupation should be defined as a set of jobs that have common requirements in terms of skills and abilities. Based on this definition, occupations should be defined in such a manner that within each group some portion of an individual's occupation-specific abilities and accumulated skills will be transferable across all jobs that fall into the group. Another important consideration is that the cost of estimating the model increases substantially as the number of occupations increases. Based on these considerations, occupations are aggregated into the five occupational groups listed in Table 1.

3.1 Descriptive Statistics

This section highlights the key characteristics of the data and provides descriptive statistics about the career choices observed in the data. Tables 2-5 show the patterns in occupational mobility for workers of different ages. Table 2 shows that there are differences in the levels of inter-firm and intra-firm occupational mobility. Mobility between occupations is more likely to occur when a person switches firms than when the person does not switch firms. The age patterns in these two types of occupational mobility are also quite different. Inter-firm occupational mobility declines by 44% from the youngest age group to the oldest, while intra-firm occupational mobility declines by 77%. The difference in the age patterns between these two types of mobility suggests that opportunities for intra-firm occupational switches may become less frequent with age.

[5] Hall (1982) provides a basis for comparison, reporting that workers, on average, hold 10 jobs over the course of their careers. Similarly, Topel and Ward (1992) find that workers hold 7 jobs in the first 10 years of their careers.

Table 3 allows for a more detailed examination of mobility between occupations. Cell (i,j) of this table (where i represents the row and j represents the column) gives the percentage of employment spells in occupation i that are followed by a spell in occupation j. For example, cell (2,1) indicates that a person employed in occupation 2 has a 7.25% chance of moving to occupation number 1, conditional on being employed in the next year. The diagonal elements of the occupational transition matrix in Table 3 are fairly large, indicating a substantial amount of persistence in occupational choices. However, even at this relatively high level of aggregation there is a substantial amount of occupational mobility. The diagonal elements show that people employed in occupation number 1 (professional, technical, and managers) are least likely to switch occupations.

Overall, the matrix is fairly symmetric, with the exception of the flows of workers between occupations 4 and 5. Workers are much more likely to move from occupation 4 (sales and clerical) and occupation 5 (service and private household) than in the opposite direction. The largest flow of workers between occupations occurs from occupation 4 (sales and clerical) to occupation 1 (professional, technical, and managers). This mobility probably reflects promotion of workers from sales into managerial positions. Mobility occurs frequently in both directions between occupations 2 (craftsmen) and 3 (operatives and laborers) and occupations 3 and 4 (service and private household).[6]

A comparsion of Tables 4 and 5 shows that occupational mobility declines steadily with age, which is consistent with human capital accumulation or job matching. Workers of all ages are most likely to remain in occupation number 1 (professional and managers). Workers between the ages of 16 and 22 who are employed in occupation number 5 (service and private household) leave the occupation over 40% of the time, but older workers ages 30-35 only leave the occupation 13% of the time. It appears that workers who are unable to move from the lowest paying occupation before they reach the final age group are very unlikely to do so.

Table 12 shows the choice distribution by age. There are 1,023 people in the sample at age 16. This number declines fairly smoothly over time because some observations are truncated at each age due to missing data. Approximately 65% of the sample attends school at age 16. Another 21% of 16 year-olds attend school and work at the same time, so the overall school attendance rate

[6] Sicherman (1990) provides an empirical analysis of occupational heirarchies in which career mobility is defined as upward mobility in a series of occupations that are ranked based on the amounts of education and training needed to work in each occupation. Ferrall (1997) estimates a structural model of hierarchies within the engineering occupation.

for 16 year-olds is 86%. School attendance takes a discrete drop to 43% at age 18, the age where most people have graduated from high school. School attendance, including attending school while working, declines steadily throughout the college ages and then drops to approximately 16% at age 22, the normal college graduation age. The data suggest that employment while attending high school is more common than employment while attending college. School attendance declines to 5% by age 25, and continues to decline at more advanced ages. Keane and Wolpin (2001) report qualitatively similar results using less highly aggregated data that divide each school year into three segments.

As school attendance declines with age, the percentage of people employed increases. A large jump in employment occurs at high school graduation, where employment increases from approximately 8% at age 17 to 30% at age 18. Employment increases to 67% by age 22, and by age 26, 85% of the people in the sample are employed and no longer attending school.

The percentage of people unemployed is 9% at age 16. Unemployment rises to 20% at ages 20 and 21 before stabilizing at close to 10% at ages 24 and above. The large number of people classified as unemployed is due to the definition of school attendance used to classify people as attending school. Recall that a person must attend school and complete a grade to be coded as attending school, so people who attend school and fail to complete a grade are classified as unemployed. Additionally, a person who is unemployed for 27 weeks during a year and employed for 25 weeks is classified as unemployed, because his primary activity during the year was unemployment. Keane and Wolpin (1997) report a similarly high rate of unemployment using slightly different definitions of employment and school attendance.

4 Economic Model of Career Choices

Each individual's career is modeled as a finite horizon, discrete time dynamic programming problem. In each year, individuals maximize the discounted sum of expected utility by choosing between working in one of the five occupations in the economy, attending school, earning a GED, or being unemployed. Employment refers only to full time work because part time workers are classified as unemployed.[7] Dual activities such as simultaneously working and attending school are also

[7] Full time work is defined as jobs where the hours worked are at least 20 per week. Including part time employment as a choice variable is conceptually straightforward but would increase the cost of estimation substantially.

feasible choices.[8] The exact set of choices available in year t depends in part on the labor force state occupied in the previous year. Each period, an individual always receives one job offer from a firm in each occupation and has the option of attending school, earning a GED, or becoming unemployed. In addition, people who are employed have the option of staying at their current job during the next year and may also have the option of switching occupations within their current firm. While employed, a worker receives either zero or one opportunity to switch occupations at his current firm.[9] Individuals observe all the components of the pecuniary and non-pecuniary rewards associated with each feasible choice in each decision period and then select the choice that provides the highest discounted expected utility.

Human capital enters the model through the endogenous accumulation of work experience and education, which affects wages and non-pecuniary utility flows. Thus, workers choose to accumulate schooling, which is costly, in order to obtain higher utility in the future. Jobs are also partly investment goods in the model because forward looking workers realize that work experience affects the distributions of wage offers and non-pecuniary benefits that they face.

4.1 Utility Function

The utility function is a choice specific function of endogenous state variables (S_t), skill endowments and preferences, and random utility shocks that vary over time, people, occupations, and firm matches. The variables in S_t measure educational attainment, firm and occupation-specific human capital, and the quality of the match between a worker and firm. To index choices for the non-work alternatives, let $s = school$, $g = GED$ and $u = unemployed$.[10] Describing working alternatives requires two indexes. Let $eq =$ "employed in occupation q", where $q = 1, ..., 5$ indexes occupations. Also, let $nf =$"working at a new firm", and $of =$"working at an old firm." Combinations of these indexes define all the feasible choices available to an individual. The description of the utility flows is simplified by defining another index that indicates whether or not a person is employed, so let $emp =$"employed". Define the binary variable $d_t(k) = 1$ if choice combination k is chosen at time t, where k is a vector that contains a feasible combination of the choice indexes. For example, $d_t(s) = 1$

[8] Light (2001) finds that omitting work experience gained while attending school produces an upward bias of 25%-44% in the estimate of the return to schooling.

[9] Many models of labor mobility ignore the possibility that workers may switch occupations within a firm. Analysis of the NLSY data presented in Section 3 suggests that that a significant fraction of workers switch occupations without switching firms.

[10] There is no uncertainty in the receipt of a GED in the model. If an individual decides to earn a GED, he receives one. In reality, people must pass a test to earn a GED. Tyler et al (2000) report that roughly 70% of people pass the GED on the first try. Within two years the eventual pass rate is 85%.

indicates that schooling was chosen at time t, and $d_t(g, e3, nf) = 1$ indicates that a GED was earned (g) while employed in the third occupation ($e3$) at a new firm (nf). Dual activities composed of combinations of any two activities are allowed subject to the logical restrictions outlined in Section 4.1.2.

4.1.1 Choice Specific Utility Flows

This section outlines the utility flows corresponding to each possible choice. The utility flow from choice combination k is the sum of the logarithm of the wage, $w_{it}^A(k)$, and non-pecuniary utility, $H_{it}(k)$, that person i receives from choice combination k at time t,

$$U_{it}(k) = w_{it}^A(k) + H_{it}(k). \tag{1}$$

The remainder of this section describes the structure of the wage and non-pecuniary utility flows in more detail.

4.1.1a *Wages.* The log-wage of worker i employed at firm j in occupation q at time t is

$$w_{it}^A = w_q(S_{it}) + \mu_i^q + \psi_{ij} + e_{ijt}. \tag{2}$$

The term $w_q(S_{it})$ represents the portion of the log wage that is a deterministic function of the work experience and education variables in the state vector. The occupation-specific subscript q allows the parameters of the wage equation to vary over occupations. For example, the effect of education on wages may differ by occupation. The term μ_i^q represents the random component of worker i's wages that is common across all firms in occupation q. This term allows people to have comparative advantages in their occupation-specific skill endowments.[11] The permanent worker-firm productivity match is represented by ψ_{ij}, and reflects match specific factors that are unobserved by the econometrician and affect the wage of worker i at firm j. True randomness in wages is captured by e_{ijt}. All of the components of the wage (w_{it}^A) are observed by the worker when a job offer is received.[12]

[11] Keane and Wolpin (2001) show that comparative advantages in occupation-specific skill endowments are an important determinant of the choice between blue and white collar employment.

[12] See Berkovec and Stern (1991) for another model where the quality of the match is revealed when drawn. In contrast, Jovanovic (1979) develops a model where agents learn about match quality over time.

4.1.1b *Non-pecuniary Utility Flows.* Non-pecuniary utility flows are composed of a deterministic function of the state vector, firm-specific match values, person specific preference heterogeneity, and random utility shocks. Define $1\{\bullet\}$ as the indicator function which is equal to one if its argument is true and equal to zero otherwise. The non-pecuniary utility flow equation is

$$\begin{aligned} H_{it}(k) \quad = \quad & [h(k, S_{it})] + \left[\phi_i^s 1\{s \in k\} + \phi_i^u 1\{u \in k\} + \textstyle\sum_{q=1}^5 \phi_i^q 1\{eq \in k\}\right] \\ & +[\varepsilon_{ikt}]. \end{aligned} \tag{3}$$

The first term in brackets represents the influence of the state vector on non-pecuniary utility flows and is discussed in more detail in the following paragraph. The second term in brackets captures the effect of person-specific heterogeneity in preferences for attending school (ϕ_i^s), being unemployed (ϕ_i^u), and being employed in occupation q (ϕ_i^q). The non-pecuniary occupation match value, ϕ_i^q, represents the random component of person i's preference for working in occupation q. It captures variation in the value that people place on job attributes such as the physical or mental demands of a job or the risk of injury that is common across jobs in each occupation. Stinebrickner (2001) shows that preference heterogeneity is an important determinant of occupational choices at the narrow level of choosing between a teaching or non-teaching job. However, this type of heterogeneity in preferences has not been extended to broader models of occupational choice. The term ϕ_i^s allows for heterogeneity in the cost of schooling caused by unobserved traits such as ability or motivation that may alter the utility cost of attending school. The final term, ε_{ikt}, is a shock to the non-pecuniary utility that person i receives from choice combination k at time t.

The remaining portion of the non-pecuniary utility function contains the non-pecuniary employment and non-employment utility flows along with the schooling cost function. This utility flow equation is specified as

$$\begin{aligned} h(k, S_{it}) \quad = \quad & \left[\textstyle\sum_{q=1}^5 \theta_q(S_{it}) 1\{eq \in k\} + \xi_{ij} 1\{emp \in k\}\right] \\ & + C^s(S_{it}) 1\{s \in k, emp \notin k\} + C^{sw}(S_{it}) 1\{s \in k, emp \in k\} \\ & + b(S_{it}) 1\{u \in k\} + C^g(S_{it}) 1\{g \in k\}. \end{aligned} \tag{4}$$

The term in brackets contains the occupation and firm-specific non-pecuniary utility flows. The occupation-specific portion of this flow, $\theta_q(S_{it})$, is a function of the state vector that is allowed to vary over occupations. This specification allows the effect of state variables such as education on employment utility to vary by occupation. The firm-specific non-pecuniary match value for person

i at firm j is represented by ξ_{ij}. This match value reflects the influence of unobservable attributes of employment at each firm that affect the employment utility flow. For example, job attributes such as commuting distance, relationships with co-workers, and availability of fringe benefits may all affect the value of a job, and their value may differ across people. Non-wage matching of this type has not been incorporated in previous models of occupational choice.[13] The second line of equation 4 contains the schooling cost function. There are two schooling cost functions, one for attending school while not employed, $C^s(S_{it})$, and one for attending school while working at the same time, $C^{sw}(S_{it})$.[14] The two schooling cost functions allow for the possibility that attending school is more costly while employed. The final components of the non-pecuniary utility flow are the deterministic portions of the value of leisure enjoyed while unemployed, $b(S_{it})$, and the cost function for earning a GED, $C^g(S_{it})$.

4.1.2 Constraints on the Choice Set

The structural modeling approach requires a detailed specification of the labor market constraints that determine an individual's choice set in each year. First, consider the case of an individual who enters time period t having not been employed in the previous year. At the start of the year the individual receives five job offers, one from a firm in each of the five occupations in the economy. Recall that a job offer consists of the wage and non-pecuniary value that the worker places on the job. The individual also observes all components of the rewards associated with attending school, earning a GED, being unemployed, and all feasible combinations of these choices.

Any dual activity is a feasible choice, subject to the following restrictions. Earning a GED must be part of a joint activity, so the single activity $d_t(g) = 1$ is not a feasible choice. In addition, earning a GED is dropped from the choice set after high school graduation or GED receipt. Finally, unemployment and employment are mutually exclusive choices. Given these restrictions, the choice set for individuals who are not employed when they enter period t is

$$\begin{aligned} D_t^{ne} \quad = \quad & \{[d_t(s), d_t(u), d_t(u,g)], [d_t(ei,nf), i = 1, ..., 5], \\ & [d_t(q, ei, nf), q = s, g,\ i = 1, ..., 5]\}. \end{aligned} \tag{5}$$

[13] Non-wage job characteristics have been shown to be an important determinant of mobility. Bartel (1982) reports that non-wage job characteristics are an important determinant of job quitting behavior. Blau (1991) rejects a reservation wage seach model in favor of a reservation utility model where hours of work affect utility.

[14] The model does not consider the effect of borrowing constraints on educational attainment. Keane and Wolpin (2001) present evidence that although borrowing constrains are severe, relaxing these constraints has little impact on educational attainment.

The first three terms correspond to the feasible non-employment opportunities, the next five terms correspond to employment in each of the five occupations, and the final ten terms are the feasible combinations of employment and attending school or earning a GED.

Next, consider the feasible choices for a person employed in occupation q. At the start of period t the individual receives one new job offer from a firm in each of the five occupations and has the option to attend school, earn a GED, or become unemployed. In addition, an employed individual always has the option of remaining at his current firm and staying in his current occupation (q). Job offers from new occupations at the current firm are received randomly, where workers receive either zero or one such offer per year. Let π_j denote the probability that a worker receives an offer to work in occupation j at his current firm, where $j \neq q$. Let π_{nq} be the probability that a worker employed in occupation q does not receive an offer to switch occupations within his current firm. This structure implies that in each period a worker always has the option of switching occupations if he switches firms, but mobility between occupations within a firm is restricted by the receipt of job offers. This feature of the model is intended to capture the fact that the scope for mobility between occupations within a firm is likely to be more limited than opportunities for mobility into new occupations when a person also switches firms.

Job offer probabilities are identified by functional form assumptions and the transition rates between occupations observed in the data. The model imposes the restriction that the distribution of the random components of job offers is the same for internal and external job offers. Given this restriction, within-firm job offer probabilities are identified by the fact that in the data, within-firm occupational switches are observed less frequently than transitions between occupations when a person moves to a new firm.[15]

The choice set for a worker employed in occupation q who receives an offer to switch to occupation j at his current firm is

$$D_t^e(j) = \{D_t^{ne}, [d_t(eq, of), d_t(s, eq, of), d_t(g, eq, of)], [d_t(ej, of), d_t(s, ej, of), d_t(g, ej, of)]\}. \quad (6)$$

If an offer to switch occupations within the current firm is not received, then the final three choices are not available to the agent. Let $D_t^e(0)$ denote this twenty-one element choice set The final restriction on the choice set is that once an individual graduates from high school or earns a GED, obtaining a GED is dropped from the choice set.

[15] See Canals and Stern (1998) for a discussion of a similar identification issue that arises in search models.

4.1.3 State Variables

The endogenous state variables in the vector S_t measure human capital and the quality of the match between the worker and his current employer. Educational attainment is summarized by the number of years of high school and college completed, hs_t and col_t, and a dummy variable indicating whether or not a GED has been earned, ged_t. Possible values of completed years of high school range from 0 to 4, and the possible values of completed college range from 0 to 5, where five years of completed college represents graduate school. Work experience is captured by the amount of firm-specific human capital (fc_t) and occupation-specific human capital (oc_t) in the occupation that the person worked in most recently. Let $O_t \in [1, 2, ..., 5]$ indicate the occupation in which a person was most recently employed. Let L_t be a variable that indicates a person's previous choice, where $L_t = \{1, ..., 5\}$ refers to working in occupations one through five, $L_t = 6$ indicates attending school full time, and $L_t = 7$ indicates unemployment.

Given this notation, the state vector is $S_t = \{hs_t, col_t, ged_t, fc_t, oc_t, O_t, L_t, \xi_t, \psi_t\}$. Including both firm and occupation-specific human capital as state variables causes problems because the size of the state space quickly becomes intractably large.[16] This is the reason that only human capital in the most recent occupation is included in the state space even though this requires a strong assumption about the transferability of human capital across occupations and the depreciation of human capital.[17] The model assumes that occupation-specific human capital is not transferable across occupations. However, age effects are included in the wage equations to proxy for general human capital that has value in more than one occupation. Human capital also depreciates completely once a person switches to a new occupation.

In addition to assuming that only human capital in the most recent occupation affects wages, a second approach is taken to further reduce the size of the state space. Assume that firm and occupation-specific human capital each take on P values, so that the possible values of human capital arranged in ascending order are

$$\begin{aligned} fc_t &\in FC = \{fc(1), ..., fc(P)\} \\ oc_t &\in OC = \{oc(1), ..., oc(P)\}. \end{aligned}$$

[16] The state space for this model contains approximately $100,000,000$ elements.

[17] Including firm specific work experience as well as experience in each of the 5 occupations increases the state space by a factor of X^6, where X is the maximum possible number of years of work experience.

After each year of work experience, with probability λ human capital increases to the next level, and with probability $(1-\lambda)$ human capital does not increase.[18] The human capital transition probability (λ) is known by agents in the model. Upon entering a new occupation, oc_t is reset to the first level. Similarly, fc_t starts at the first level in the first year of employment at a firm. When the maximum level of human capital is attained, no further increases are possible. The size of the state space is significantly reduced when P is a small number relative to the possible values of work experience, but the model still captures the human capital improvement process. In this work, $P = 3$.

This method of modelling human capital has the advantage of making it possible to include both firm and occupation-specific human capital in the state space at a fraction of the cost of keeping track of actual years of experience at a firm or in an occupation, because work experience could range from zero to fifteen years in this model. In models of this type with large state spaces, an alternative approach would be to place relatively low upper bounds on state variables, or omit some of them entirely. The approach presented here is appealing from a practical standpoint because it makes estimation feasible, but it is also consistent with the theory of human capital. The number of years of completed work experience is generally included as an explanatory variable in wage regressions only as a proxy for the unobservable level of human capital that actually affects wages. Viewing increases in human capital as a stochastic event is consistent with this idea, because it allows for the possibility that years of work experience may vary for people with a given level of human capital.

4.2 The Optimization Problem

Individuals maximize the present discounted value of expected lifetime utility from age 16 ($t = 1$) to a known terminal age, $t = T^{**}$. At the start of his career, the individual knows the human capital wage function in each occupation, as well as the deterministic components of the utility function. An individual also knows his endowment of market skills (μ's) and occupation-specific non-pecuniary match values (ϕ's). Future realizations of firm-specific match values (ψ's and ξ's) and time and choice specific utility shocks (ε's and e's) are unknown. Although future values are unknown, individuals know the distributions of these random components. Individuals also know their current levels of firm and occupation-specific human capital (fc_t and oc_t), as well as the

[18] Brown and Flinn (2002) use a similar method to model the process by which child quality changes over time.

probability that human capital will increase in the next period, conditional on employment (λ).

The maximization problem can be represented in terms of alternative specific value functions which give the lifetime discounted value of each choice for a given set of state variables, S_t. Variation in the structure of the value functions comes from differences in the utility flows across states, and differences in the choice set across states. Regarding the choice set, there are only two relevant categories of states: employed (including joint employment activities), and all other choice combinations. While people are employed, the possibility of mobility between occupations within their current firm implies that the value function will be structured differently than when non-employed, because the employed value function must incorporate the value of internal job offers.

The value function for an individual with discount factor δ employed in occupation q is the utility flow from employment, plus the expected value of the best choice available next period,

$$V_t(eq, l) = U_t(eq, l) + \delta \sum_{k \neq q} \pi_k EZ_t^{ek} + \delta[\pi_{nq} EZ_t^{eq}], \qquad q = 1, ..., 5, \quad l = of, nf. \tag{7}$$

The EZ_t^{ek} terms represent the expected value of the best choice in period $t+1$, conditional on receipt of an offer to work in occupation k at the worker's current firm. The expectations are taken over the random components of the choice specific utility flows, which are the random utility shocks and match values, $\{\varepsilon, e, \psi, \xi\}$. The expectation is also taken over firm and occupation-specific human capital, (fc and oc) since human capital evolves stochastically.[19]

Consider the first summation in equation 7. Each term in the sum corresponds to the probability that a job offer to work in a new occupation at the current firm is received (so $k \neq q$), multiplied by the corresponding expected value of the best option next period. For each occupation q it must be the case that $\sum_{j \neq q} \pi_j + \pi_{nq} = 1$. The structure of the value function is similar to the model presented by Wolpin (1992) in that both models allow the arrival of some types of job offers to be random, which implies that the values of future choices must be weighted by job offer probabilities. Wolpin (1992) estimates job offer probabilities for unemployed and employed job searchers, in contrast to the intra-firm job offer probabilities estimated in the present model.

The individual elements of the EZ_t^{ek} terms are the time $t+1$ value functions for each feasible

[19] See Rust and Phelan (1997) for an example of another dynamic programming model where agents face uncertainty about how the state vector will evolve over time.

choice,

$$\begin{aligned} EZ_t^{ek} &= E\max\{V_{t+1}(s), V_{t+1}(u), V_{t+1}(u,g), [V_{t+1}(ei,nf), V_{t+1}(m,ei,nf), \\ &\quad m=s,g,\ i=1,...,5,], V_{t+1}(eq,of), V_{t+1}(s,eq,of), V_{t+1}(g,eq,of), \\ &\quad V_{t+1}(ek,of), V_{t+1}(s,ek,of), V_{t+1}(g,ek,of)\}. \end{aligned} \tag{8}$$

In the remainder of the paper, I will refer to these expected values as "Emax". The final term in the employed value function corresponds to the case where an individual does not receive an offer to switch occupations within his current firm. In this case, switching occupations without switching firms is not possible, so the expected value of the best choice at time $t+1$ is

$$\begin{aligned} EZ_t^{eq} &= E\max\{V_{t+1}(s), V_{t+1}(u), V_{t+1}(u,g), \\ &\quad [V_{t+1}(ei,nf), V_{t+1}(m,ei,nf), m=s,g,\ i=1,...,5], \\ &\quad V_{t+1}(eq,of), V_{t+1}(s,eq,of), V_{t+1}(g,eq,of)\}. \end{aligned} \tag{9}$$

The value function for an individual who is not currently employed is simpler because mobility within a firm is obviously not possible for people who are not employed. The value function is

$$V_t(p) = U_t(p) + \delta EZ_t^{su}, \qquad p = s,u \tag{10}$$

$$V_t(u,g) = U_t(u,g) + \delta EZ_t^{su}. \tag{11}$$

The corresponding expected value of the maximum term is

$$\begin{aligned} EZ_t^{su} &= E\max\{V_{t+1}(s), V_{t+1}(u), V_{t+1}(u,g), \\ &\quad V_{t+1}(ei,nf), V_{t+1}(m,ei,nf),\ m=s,g,\ i=1,...,5\}, \end{aligned} \tag{12}$$

which consists of all feasible combinations of schooling, unemployment, and new job offers.

Agents making career decisions use the value functions to determine the optimal educational and employment choices in each period. Each period, a person observes all of the components of the utility flows of each feasible choice, and then calculates the value of each choice using equations 7 through 12. He then chooses the option with the highest discounted expected value.

4.3 Solving the Career Decision Problem

Estimating the structural parameters of the model requires solving the optimization problem faced by agents in the model. The finite horizon dynamic programming problem is solved by backwards recursion. Assume that there is some age, T^*, after which no choices are made, and another age, T^{**} at which the agent dies. Then, evaluating the value functions from T^* to T^{**} is straightforward, because the value function for each choice is simply a sum of one period expected utility flows. Given the value functions at age T^*, the value functions can be solved backwards recursively for all $t < T^*$ using equations 7 through 12. Before considering the solution of the model in more detail, it is useful to specify the distributions of the random components of utility flows.

4.3.1 Distributional Assumptions

Assume that firm-specific match values and randomness in wages are distributed i.i.d normal,

$$\begin{aligned} \xi_{ij} &\backsim N(0, \sigma_\xi^2) \\ \psi_{ij} &\backsim N(0, \sigma_\psi^2) \\ e_{ijt} &\backsim N(0, \sigma_e^2). \end{aligned}$$

The firm-specific pecuniary and non-pecuniary match values are part of the state space because the value function associated with a job depends on the wage match value (ψ_{ij}) and non-wage match value (ξ_{ij}) for worker i at firm j. The distributions of these variables are continuous, which causes a problem because the state space becomes infinitely large when continuous variables are included. This problem is solved by using a discrete approximation to the distributions of wage match values (ψ_{ij}) and non-wage match values (ξ_{ij}) when solving the value functions and computing the likelihood function.[20]

Assume that the random choice-specific utility shocks are distributed extreme value, with distribution function

$$F(\varepsilon) = \exp\{-\exp(-\frac{\varepsilon}{\tau})\},$$

and with variance $\pi^2\tau^2/6$. The assumption that the ε's are distributed extreme value simplifies the computation of the value functions and choice probabilities.[21]

It remains to specify the distributions of the occupation-specific skill endowments (μ's) and preferences (ϕ's). Using an approach similar to Heckman and Singer (1985), Keane and Wolpin (1997), and Stinebrickner (2001), the joint distribution of skill endowments and preferences is

[20] Currently, the discrete approximations to ξ and ψ have three points of support each. Increasing the number of points in the approximation would increase the computational burden of estimation.

[21] See Rust (1987) and Berkovec and Stern (1991) for examples.

specified as a discrete multinomial distribution. Let $\Phi_i = \{\mu_i^1, ..., \mu_i^5, \phi_i^1, ..., \phi_i^5, \phi_i^s, \phi_i^u\}$ be the vector of skill endowments and preferences that are known to the agent at age sixteen.

Assume that there are M types of people, each with a different endowment of skills and preferences, $\{\Phi_m, m = 1, ..., M\}$. Define χ_m as the proportion of the mth type in the population. Endowment heterogeneity is unobserved to the econometrician, but assume that we do know that there are M types of people. This flexible assumption about the joint distribution of skills and preferences allows for a wide range of patterns of comparative advantages in skills and heterogeneity in preferences. As the number of types of people, M, becomes large, this approach can approximate any joint distribution of skills and preferences arbitrarily well.

4.3.2 Calculating the Value Functions

This section discusses the details of the solution to the dynamic programming problem. The major complication arises from the fact that as the model is specified the Emax integrals do not have closed form solutions. In many dynamic programming models, researchers assume that the only randomness in utility flows is choice specific, independent over time, and distributed extreme value. See Rust (1997) for a recent example. A consequence of this assumption is that the Emax integrals have a simple closed form. However, the unappealing consequence of this assumption in this application is that it rules out job matching in wages and non-wage utility flows. Job matching in wages has been shown to be empirically important in papers such as Miller (1984) and Berkovec and Stern (1991).

To the extent that mobility decisions are based on non-wage factors, the addition of matching in non-wage utility flows to the career decision problem will contribute to the understanding of the causes of transitions between firms and occupations. This work allows for job matching effects by using simulation methods to evaluate the high dimensional integrals required to calculate Emax. Berkovec and Stern (1991) avoid having to use simulation methods because they assume that people know their future job match values with certainty. Allowing for uncertainty in future realizations of job match values provides a more complete description of the factors influencing mobility between firms.

At this point it is useful to partition the vector of error terms, excluding ε, into two sets. Let $\Omega_t = \{\psi, \xi, e\}$ be the set of errors whose future realizations are unknown to the agent at time t, and define the joint density of these errors as $f(\Omega_t)$. Recall that the vector of skill endowments and preferences is $\Phi_i = \{\mu_i^1, ..., \mu_i^5, \phi_i^1, ..., \phi_i^5, \phi_i^s, \phi_i^u\}$. Consider calculating the expected value of the best choice available next period for a person who is employed in the current time period. Conditional on Ω_t and firm and occupation-specific human capital (fc_t and oc_t), the expected value of the maximum has a closed form solution because of the assumption that ε is distributed extreme value,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \mid \Omega_t, \Phi_i, oc_t, fc_t\} &= \tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t \mid \Omega_t, \Phi_i, oc_t, fc_t)}{\tau})]) \\ &= \Psi(d_t \,| \Omega_t, \Phi_i, oc_t, fc_t) \;, \end{aligned} \tag{13}$$

where $\bar{V}(d_t) = V(d_t) - \varepsilon$, γ is Euler's constant, and τ is a parameter of the extreme value distribution. Let $f(\bullet)$ represent the density of the variable in parentheses. Integrating over the distributions of Ω_t, fc_t and oc_t provides the unconditional expected value of the best choice available next period,

$$E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \mid \Phi_i\} = \int \int \left[\int \cdots \int \Psi(d_t \,| \Omega_t, \Phi_i, oc_t, fc_t) f(\Omega_t) d\Omega_t \right] f(fc_t) dfc_t f(oc_t) doc_t. \tag{14}$$

This integral does not have an analytic solution, so it is simulated using R draws from the joint density $f(\Omega_t)$. In this work, $R = 20$.[22] The integral over the distribution of human capital is simply a probability weighted sum because the distribution of human capital is discrete. Let r index simulation draws, and the simulated integral is simply the average of equation 14 over the R draws,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon | \; \Phi_i\} &= \frac{1}{R} \sum_{r=1}^{R} \sum_{h=1}^{P} \Pr[fc_t = fc_t(h) \,|\, fc_{t-1}] \sum_{z=1}^{P} \Pr[oc_t = oc_t(z) \,|\, oc_{t-1}] \times \\ & \quad \Psi(d_t \,| \Omega_t^r, \Phi_i, oc_t^z, fc_t^h). \end{aligned} \tag{15}$$

The other Emax terms found in the value function calculations are approximated using this method.

The major computational burden of solving the model arises from the fact that the Emax functions must be simulated at each point in the state space over the agent's entire time horizon. When the number of points in the state space is large, as it is in this model, evaluating the value function becomes very time consuming. Several methods to reduce the computational expense of evaluating value functions in dynamic programming models have been developed in recent years. For example, Rust (1997) proposes a method that uses randomization to break the curse of dimensionality, Keane and Wolpin (1994) use a linear regression to interpolate value functions, and Brien, Lillard,

[22] Antithetic acceleration is used throughout estimation to reduce variance of the simulated integrals. See Geweke (1988) for a discussion of antithetic acceleration, and Stern (1997) for a review of the applications of simulation methods in the economics literature.

and Stern (2003) interpolate value functions using a weighted average of "close" points in the state space.

This paper employs an interpolation algorithm that follows along the lines of the one developed by Keane and Wolpin (1994). As in Keane and Wolpin (1994), value functions are simulated at a fraction of the state space and interpolated at the remaining points in the state space. This paper implements a new regression function that exploits the assumption that ε is distributed extreme value. If the only source of randomness in the model was the error term ε, then the expected value of the maximum would have the closed form solution shown in equation 13. This is not the case in this model due to the existence of the wage match values (ψ), non-wage match values (ξ), and random wage shocks (e), but it suggests the following functional form for the interpolating regression

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon\ \} &= \omega_{0t} + \omega_{1t}\tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t)}{\tau})]) \\ &= \omega_{0t} + \omega_{1t}\Psi(d_t)\ . \end{aligned} \tag{16}$$

The parameters ω_{0t} and ω_{1t} are estimated by OLS, and allowed to vary over time. This regression function has the desireable theoretical property that it converges to the exact solution for Emax as σ_ξ, σ_ψ, and σ_e approach 0. In addition, it also satisfies the theoretical restrictions on the Emax function outlined in Stern (1991). Another important property of this regression function is that the regressor is defined at every point in the state space even if set of feasible choices varies over points in the state space, as it does in this model. In contrast, the regression function proposed by Keane and Wolpin (1994) uses the value functions corresponding to each element in the choice set separately as regressors, which creates a missing data problem when the choice set is state dependant.[23]

During estimation, the value functions are simulated at approximately 1% of the state space and interpolated at the remaining points. The regression function fits the data very well. Throughout estimation, the R^2 from the interpolating regression remained between .98 and .99. Experimentation shows that the actual and interpolated value functions differ by approximately 1% on average.

[23] One solution to this problem would be to use a different interpolating regression for each feasible choice set in the state space.

5 Estimation of The Structural Model

The parameters of the model are estimated by maximum simulated likelihood (MSL) using the career history data from the NLSY. This section begins by specifying functional forms for the utility flow equations. It concludes with a derivation of the likelihood function and a discussion of the methods used to maximize the likelihood function.

5.1 Further Model Specification

Before discussing the details of estimating the parameters of the structural model, it remains to specify the wage equations, non-pecuniary utility flow equations, and job offer probabilities in more detail.

5.1.1 Utility Flow Equations

This section defines the deterministic portion of the utility function. The deterministic portion of the occupation-specific human capital wage function is

$$\begin{aligned} w_q(S_{it}) \quad = \quad & \beta_1^q age_{it} + \beta_2^q hs_{it} + \beta_3^q col_{it} + \beta_4^q 1[age_{it} \leq 17] + \\ & \beta_5^q 1[age_{it} \geq 18 \cap age_{it} \leq 21] + \beta_6^q ged_{it} \\ & + \beta_7^q 1[fc_{it} = fc(1)] + \beta_8^q 1[fc_{it} = fc(2)] + \beta_9^q 1[fc_{it} = fc(3)] \\ & + \beta_{10}^q 1[oc_{it} = oc(1)] + \beta_{11}^q 1[oc_{it} = oc(2)] + \beta_{12}^q 1[oc_{it} = oc(3)]. \end{aligned} \tag{17}$$

The parameters β_7^q and β_{10}^q are set equal to zero because the first level of human capital is not separately identified from the constant in the wage equation (μ's). The explanatory variable age_t is included as a proxy for general human capital that is transferable across all firms and occupations.

Let NF_t be a dummy variable indicating whether or not the individual is in his first year of employment at a firm after being employed at a different firm in the previous period. Let hd_t and cd_t represent dummy variables that indicate receipt of a high school or college diploma. The non-pecuniary utility flow equation for occupation q is

$$\begin{aligned} \theta_q(S_{it}) \quad = \quad & \alpha_1 age_{it} + \alpha_2 hs_{it} + \alpha_3 col_{it} + \alpha_4 hd_{it} + \alpha_5 cd_{it} + \alpha_6 ged_{it} + \alpha_7 fc_{it} + \alpha_8 oc_{it} \\ & + \alpha_9 1[L_{it} > 5] + \alpha_{10} NF_{it} \qquad\qquad q = 1, ..., 5 \end{aligned} \tag{18}$$

The parameters of the non-pecuniary occupation-specific utility flows are constrained to be the same across all occupations to reduce the number of parameters that must be estimated.

The cost function for attending school is

$$\begin{aligned} c^S(S_{it}) &= \gamma_{s1} age_{it} + \gamma_{s2} hd_{it} + \gamma_{s3} cd_{it} + \gamma_{s4} hs_{it} + \gamma_{s5} col_{it} + \gamma_{s6} 1[L_{it} \neq 6] \\ c^{SW}(S_{it}) &= \gamma_{sw1} age_{it} + \gamma_{sw2} hd_{it} + \gamma_{sw3} cd_{it} + \gamma_{sw4} hs_{it} + \gamma_{sw5} col_{it} + \gamma_{s6} 1[L_{it} \neq 6] \end{aligned} \tag{19}$$

The data do not contain information about the monetary cost of attending school, so it is not possible to separately identify the pecuniary and non-pecuniary cost of attending school. This limitation implies that the schooling utility flow represents the non-pecuniary benefit of attending school minus the pecuniary and non-pecuniary costs. The deterministic portion of the unemployment utility flow, $b(S_{it})$, is set equal to zero because the non-wage utility flow coefficients are only identified relative to a base choice, as in any discrete choice model.

The final utility flow equation represents the utility derived from earning a GED. The deterministic portion of the GED utility flow is

$$c^g(S_{it}) = \gamma_{g1} + \gamma_{g2} age_{it}. \tag{20}$$

Within-firm job offer probabilities are specified as multinomial logit, so the probability of receiving a job offer from occupation j at the current firm is

$$\pi_j = \frac{\exp(\rho_j)}{\sum_{k=1}^{5} \exp(\rho_k)}. \tag{21}$$

Finally, the discount factor, δ, is set equal to .95 rather than estimated because it can be difficult to estimate the discount factor in dynamic models, even though it is technically identified.[24]

5.1.2 Measurement Error

Empirical evidence suggests that wages are measured with error. Let w^o represent the logarithm of the wage that is observed in the data, and let w^A represent the logarithm of the true wage. Suppose that the observed wage is equal to the true wage plus a noise term, ζ. The relationship between the actual and observed wage is

$$w^o = w^A + \zeta, \tag{22}$$

[24] See Berkovec and Stern (1991) for an example of a model where it was not possible to estimate the discount factor. Rust and Phelan (1997) find that the likelihood function for their dynamic retirement model is very flat as a function of the discount factor, so they estimate the discount factor using a grid search. Keane and Wolpin (1997) are able to estimate a yearly discount factor, finding that it is .936.

where $\zeta \sim N(0, \sigma_\zeta)$. The solution to the dynamic programming problem is unchanged by the addition of measurement error because agents make decisions based on their true wages, not the wages observed in the data. However, the likelihood function must be modified to account for the fact that the observed and actual wages may differ.

5.2 The Likelihood Function

The likelihood function used to estimate the structural model follows directly from the model presented in Section 4. The solution to the dynamic programming problem provides the choice specific value functions which are used in the construction of the likelihood function. The vector of parameters, denoted by Θ, is made up of the parameters found in the deterministic portions of the choice-specific utility flows, error standard deviations, job offer probabilities, and skill endowment vectors and type probabilities. Define O_{it} as the observed outcome for person i at time t, which consists of an observed choice and possibly an observed wage. The likelihood contribution for person i at time t is simply the joint probability of the choice made by the person and the wage, if one is observed. These probabilities are discussed in more detail below.

Conditional on having an endowment vector of type k, the likelihood contribution for person i is the product of the probabilities of each outcome observed in the data over the $\widetilde{T}_i$ years that the person remains in the sample, conditional on the observed state variables,

$$\begin{aligned} L_i(\Theta \mid \Phi_i = \Phi_k) \quad = \quad & \int \cdots \int [\int \int \left(\prod_{t=1}^{\widetilde{T}_i} \int \Pr[O_{it} \mid \Theta, S_{it}, \Phi_i = \Phi_k) dF(\zeta) \right) \\ & dF(oc) dF(fc)] dF(\Omega). \end{aligned} \tag{23}$$

Note that the path probability for each person is integrated over the distributions of occupation and firm-specific human capital (oc and fc) because these variables are unobserved. The likelihood contribution is also integrated over the joint distribution of Ω, because these match values and choice specific utility shocks are not observed.

This high dimensional integral is simulated using R draws from the joint distribution of Ω and Q draws from the joint distribution of occupation and firm-specific human capital. The integral over the joint distribution of human capital is simulated using a modified Geweke, Keane, and Hajivassiliou (GHK) algorithm because the joint distribution of human capital is intractably complex. The details of this algorithm are provided in Appendix A. The innermost integral over the distribution of measurement error is simulated using B draws from the distribution of ζ. The simulated likelihood contribution is

$$
\begin{aligned}
L_i^R(\Theta \mid \Phi_i = \Phi_k) \quad = \quad & \frac{1}{R}\sum_{r=1}^{R}\frac{1}{Q}\sum_{q=1}^{Q}\prod_{t=1}^{\widetilde{T}_i}\frac{1}{B}\sum_{b=1}^{B} \\
& \Pr[O_{it}^{rqb} \mid \Omega_i^r, \zeta^b,\, oc^q, fc^q, \Theta, S_{it}, \Phi_i = \Phi_k).
\end{aligned}
\tag{24}
$$

The unconditional simulated likelihood contribution is a weighted average of the type-specific likelihoods, where the weights are the type probabilities,

$$L_i^R(\Theta\,) = \sum_{m=1}^{M}\chi_m L_i^R(\Theta \mid \Phi_i = \Phi_m). \tag{25}$$

The likelihood function for the entire sample is simply the product of the likelihood contributions for each person,

$$L^R(\Theta\,) = \prod_{i=1}^{N} L_i^R(\Theta\,). \tag{26}$$

The vector of parameters $\widehat{\Theta}$ that maximizes equation neber 26 is the MSL estimate of Θ. The result of the MSL estimation routine, $\widehat{\Theta}$, is a consistent estimator of Θ as the number of simulation draws tends to infinity. However, Borsch-Supan and Hajivassiliou (1993) provide Monte Carlo evidence that MSL provides precise parameter estimates even for a small, fixed number of simulation draws. The likelihood function is simulated using 20 draws from the joint distribution of firm and occupation specific human capital, and 20 draws from the joint distribution of the errors. Antithetic acceleration is used to reduce the variance of the simulated integrals.

5.2.1 Outcome Probabilities

The most straightforward outcome probability found in the likelihood function is the probability of observing a person attending school or being unemployed. In order to make things concrete, consider the likelihood contribution for a person attending school in time t who was not employed in period $t-1$. The likelihood contribution is simply the probability that the value of attending school exceeds the value of any other choice in the person's choice set, D_t^{ne}. A consequence of the assumption that ε is distributed extreme value is that conditional on the other error terms (Ω), endowment vector (Φ_i), and occupation and firm-specific human capital (oc and fc), the choice probability is of the multinomial logit form,

$$\Pr(d_{it} = s \mid \Omega,\, oc, fc, \Theta, S_{it}, \Phi_i) = \frac{\exp(V_t(s))}{\sum\limits_{k\in D_t^{ne}}\exp(V_t(k))}. \tag{27}$$

The numerator contains the value of attending school in period t, and the denominator contains the value functions for each of the feasible choices at time t. Computing the unconditional likelihood contribution requires integrating over Ω, oc, and fc as discussed previously.

The probabilities for outcomes involving employment are similar to the non-employed outcome probabilities, except they also include the density of the observed wage. The probability of interest is the joint probability of the observed choice (k), and the accepted wage observed in the data (w^o),

$$\begin{aligned}\Pr(d_{it}(k) &= 1 \cap w = w^o) = \Pr(d_{it}(k) = 1 \,|\, w = w^o) \Pr(w = w^o) \\ &= \Pr(d_{it}(k) = 1 \,|\, w = w^o) \Pr(w^A + \zeta = w^o)\end{aligned} \tag{28}$$

For example, consider the probability of observing a person taking a job in period t with an observed wage of wage of w^o at a new firm in occupation number one when the person was not employed in period $t-1$. Let $f_e(\bullet)$ denote the density of the wage error, e. The outcome probability is

$$\begin{aligned}\Pr(d_{it}(e1, nf) &= 1 \;\cap\; w = w^o \,|\, \Omega, oc, fc, \Theta, S_{it}, \Phi_i) = \\ &\frac{\exp(V_t(e1, nf \,|\, e_{1t} = w^o - w^A_{i1t} - \zeta))}{\sum\limits_{k \in D^{ne}_t} \exp(V_t(k \,|\, e_{1t} = w^o - w^A_{i1t} - \zeta))} \\ &\times f_e(w^o - w^A_{i1t} - \zeta)\;.\end{aligned} \tag{29}$$

Again, this probability must be integrated over the error terms in Ω and over the joint distribution of human capital to produce the unconditional likelihood contribution. When the accepted wage of a job is not observed due to missing data, the wage density is not part of the outcome probability and the choice probability is not conditioned on the observed wage.

5.3 The Estimation Algorithm

The parameters of the structural model are estimated using a derivative based optimization routine. Starting from an initial guess of the parameters, derivatives of the likelihood function are used to update $\widehat{\Theta}$. This iterative process is computationally expensive because the dynamic programming problem must be solved at each iteration of the parameter vector in order to calculate the likelihood function. Also, computing numerical derivatives requires solving the dynamic programming problem again for each parameter in the vector Θ. The covariance matrix of $\widehat{\Theta}$ is estimated using the "outer product of the gradient" method of Berndt, Hall, Hall, and Hausman (1974).

The computational burden of estimation is reduced by using parallel processing while evaluating the derivatives. See Swann (2001) for a user friendly discussion of how to apply parallel processing to a simple maximum likelihood problem. The techniques employed in this paper are slightly different from those discussed by Swann, but rely on the same parallel processing techniques. Suppose there are P parameters in the model, which means that P costly numerical derivatives must be computed at each iteration of the estimation algorithm. In the parallel program these P derivatives are divided evenly among C available processors, which decreases computation time by a factor of C.[25] There is some additional cost to employing parallel processing techniques relative to a serial program, because it takes time to send messages between processors. This cost is minimal in this application, so the program runs very close to C times faster than a serial version.

6 Structural Parameter Estimates

Table 7 presents the preliminary structural parameter estimates and the associated standard errors. There are too many parameters (135 to be exact) to discuss each one individually, so instead the following discussion focuses on key coefficient estimates and what the estimates reveal about the career decision process.

6.0.1 The Log Wage Equation

The estimates of the parameters in the log wage equation reveal the importance of occupation and firm-specific human capital in determining wages. The estimates of the human capital levels are statistically significant in all occupations, and the magnitudes of the levels are large relative to the other parameters in the wage equation. The first human capital level in each occupation is fixed at zero because these parameters are not separately identified from the constant in the wage equation. The probability that human capital increases after each year of tenure is fixed at .15 for each occupation. In future versions of the model this parameter will be estimated to allow the rate of skill increase to vary across occupations. The firm and human capital parameter estimates reveal the percentage change in wages that a worker experiences when he moves to a higher level of human capital. For example, the second level of firm specific human capital in the professional and

[25] In this work, the number of processors used ranged from 8 to 30, due to constraints on available computing resources.

managerial occupation is .09, which means that moving from the first to second level of firm-specific human capital increases wages by 9%. Of course, this does not mean that the estimate of the yearly tenure effect is 9% per year, because there is only a 15% chance that a person's skills will increase after one year of tenure.

A more meaningful way to examine the returns to human capital is to examine the distribution of wage increases due to human capital accumulation at different levels of tenure. This type of analysis takes into account both the human capital levels and the rate of increase in skills. Table 8 presents the mean and variance of the returns to firm and occupation-specific human capital for workers with 2, 6, and 10 years of continuous tenure in an occupation or at a firm. For example, the first entry in the table indicates that 2 years of firm tenure increase wages by an average of 3.9% for a professional and managerial worker. The standard deviation of this percentage change is .111. The average wage increases attributed to the accumulation of human capital suggest that both firm and occupation specific human capital play important roles in determining wages. The wage gains associated with firm and occupation-specific human capital vary widely across occupations. For example, the average return to 6 years of tenure at a firm ranges from a low of 9.6% for operatives and laborers to a high of 39% for sales and clerical workers. The average return to six years of tenure in an occupation ranges from 7.4% for craftsmen to a high of 59% for service workers.

There is still disagreement in the literature over the effects of firm and occupation specific capital on wages. For example, estimates of the return to 10 years of tenure at a firm range from Altonji and Shakotko's (1987) estimate of 6.6% to Topel's (1991) estimate of over 25%. Altonji and Shakotko use deviations from mean tenure on a job spell as an instrument for the endogenous job tenure, and Topel uses a closely related two stage estimation procedure. Comparing the results of these studies to the results reported in this paper is complicated by the fact that all published studies of the tenure effect ignore the possibility that the return to firm tenure may vary over occupations. The estimates of the effect of 10 years of firm tenure on wages reported in this paper range from 12% to 50% across occupations. Overall, these estimates are at the higher end of those reported in the literature.

The returns to occupation specific-human capital have only been examined in one other study. Kambourov and Manovskii (2002) report that 10 years of occupational tenure increase wages by approximately 19%, although they do not allow this effect to vary across occupations. The estimates reported in this paper range from 15% to 121%, so they are higher than the Kambourov and

Manovskii estimates.

The relative importance of firm and occupation-specific human capital varies substantially across occupations. For example, the returns to firm and occupational tenure are approximately the same for professional and managerial workers. In contrast, for operatives and laborers the returns to occupational tenure are more than twice as large as the returns to firm tenure. The standard deviations of the returns to tenure are also quite large, which indicates that there is quite a bit of variation in the returns to human capital for workers with a given number of years of tenure.

Returning to Table 7, the estimates of the log wage equation provide information about the effects of age and education on wages for each of the five occupational groups. The parameter estimates show that wages decrease with age in each occupation, after controlling for firm and occupation specific human capital. The age effect is nonlinear, with workers age 21 and younger earning substantially less than workers over the age of 21.

The estimated effects of high school and college on wages vary across the five occupations. Attending high school has a negative effect on wages for professional workers, craftsmen, and service workers. In contrast, the effect of a year of college on wages is positive in each occupation. The effect of a year of college on professional wages is quite large, at 19%. If a person completes both high school and college, their wages as a professional worker or manager will be 60% higher than someone with zero years of completed education ($-.039 \times 4 + .19 \times 4 = .60$). In contrast, a service worker with the same amount of education will realize a wage gain of only 15.6% ($-.033 \times 4 + .072 \times 4 = .156$). The variation in returns to schooling across occupations implies that a person's educational choices will be influenced by their expectations about their future occupational choices.

The estimates of the effect of a GED on wages are statistically significant in each of the five occupations. The estimates range from $-.497$ to $.743$ across the five occupations. In contrast, Cameron and Heckman (1993) find that the GED does not have a significant effect on wages using a regression which assumes that earning a GED is exogenous. In contrast, Tyler, Murnane, and Willett (2000) use a natural experiment approach based on variation in the GED passing standard across states to determine that the GED increases wage by 10-19%.

6.0.2 Non-pecuniary Utility Flows

The non-pecuniary utility flow coefficient estimates are presented in the second page of Table 7. These parameter estimates represent the effect of each variable on the non-pecuniary utility derived from attending school, earning a GED, or being employed relative to the base choice of unemployment. For example, one more year of age decreases the value of attending school (relative to unemployment) by 2.356 log yearly wage units. The decrease in the schooling utility flow caused by a one year increase in age is equivalent to the effect of a 91% decrease in the yearly wage on employment utility, so the decline in schooling utility with age is quite large in magnitude. The estimates of the college and graduate school attendance dummy variables show that the consumption value of college is higher than the consumption value of high school. Utility also declines sharply in graduate school relative to the undergraduate years.

The model also allows observable characteristics to affect non-pecuniary employment utility. These coefficient estimates show that age increases the non-pecuniary utility flow from employment. In contrast, education decreases the non-pecuniary employment utility flow. Earning a high school diploma or GED also increases employment utility. The non-pecuniary employment utility flow increases with both occupation and firm-specific human capital.

6.0.3 Labor Market Constraints: Switching Costs & Job Offer Probabilities

The magnitude and statistical significance of the switching costs indicates that there are substantial non-pecuniary costs of moving between jobs and of moving to a new job after a period of non-employment. Incurring the firm switching cost of 3.769 log yearly wage units has the same effect on one period utility as a 98% decrease in wages. The cost of moving from unemployment to employment is 2.775, which has the same affect on utility as a 94% decrease in wages. Although workers must pay a large one time non-pecuniary cost when starting a new job, workers in the dynamic model realize that the discounted benefits in terms of higher wages and utility are received over the duration of the new job. The estimates also indicate that a worker's utility is increased by .869 when a worker returns to school after being employed in the previous year.

The estimates also show that there are substantial non-pecuniary costs associated with attending school while working. Not surprisingly, the largest cost is associated with working while attending graduate school. Working full-time reduces the amount of time that a person has available for studying and extracurricular activities, which would tend to increase the cost of attending school and decrease the consumption value of attending school.

The estimates of the within-firm job offer probabilities show that people are least likely to receive internal job offers to become service workers. Workers are most likely to receive a job offer to become professional or managerial workers while remaining with their current employer. Overall, the job offer probabilities suggest that intra-firm occupational mobility is restricted by the arrival of job offers.

6.0.4 Heterogeneity in Skills and Preferences, Job Matching, and Randomness

The final section of Table 7 presents the log-wage equation intercepts and non-pecuniary utility flow intercepts for the three types of people in the model, along with the estimated proportion of each type in the population. The log wage intercepts represent skill endowments in each of the five occupations, and the non-wage intercepts reflect preferences for employment in each occupation and for attending school. The estimates of the log wage intercepts show that a fairly complex pattern of comparative advantage is present in the sample. Type #1's, the largest group in the population, are the most productive as professionals and managers or sales and clerical workers, but are the least productive when employed as craftsmen or service workers. Type #2's are ranked second as both craftsmen and service workers, and ranked last in the other three occupations. Type #3's have the highest productivity when working as craftsmen, operatives and laborers, or service workers.

There are fairly large differences in the endowments of occupation-specific skills across the three types of people. Differences in the log wage intercepts correspond to percentage changes in wages, so a person's endowment type greatly influences their expected earnings in each occupation. For example, holding all state variables constant, a type #1 person's expected wage in the sales and clerical occupation is 75% higher than a type #2 person's expected wage, and 30% higher than a type #3 person's expected wage as a sales and clerical worker.

The non-pecuniary intercepts reflect a person's preferences for working in each occupation and attending school. These parameters are measured in log yearly wage units relative to the base choice of unemployment. The preference for attending school (or school ability) represents the consumption value of school, net of the pecuniary and non-pecuniary costs of attending school. The value of attending school varies substantially across types, from a low of 5.19 log yearly wage units for type #1's, to a high of 13.41 for type #2's. To get a feel for the magnitude of the consumption value of attending school, it is useful to calculate the consumption value of school attendance in dollars relative to the value of being unemployed for a 16 year-old who has completed two years of high school. The value relative to unemployment is $6.72 for a type #1 person, $27,363 for a type #2 person, and \$1,246 for a type #3 person. Averaged over types, the consumption value of attending school for this typical 16 year-old is \$9,394 greater than the value of being unemployed. While preferences for attending school vary substantially across types, it is important to remember that all types of people may choose to invest in education, because education has value as an investment good as well as a consumption good.

The non-pecuniary employment utility intercepts vary substantially across occupations and types. Type #3's place a higher value on working in each of the five occupations relative to the value of being unemployed. The magnitudes of the differences in the non-pecuniary intercepts for each type of person suggest that preferences play a large role in determining occupational choices. For example, the effect of moving from the craftsmen occupation to the professional occupation varies substantially across the three types of people, and is large in magnitude. The change in non-pecuniary utility accompanying a switch from the craftsmen to the professional occupation is -1.647 utils for a type #1 person, .693 utils for a type #2 person, and $-.944$ utils for a type #3 person. These non-pecuniary utility changes are equivalent to an 80% decrease in wages for a type #1 person, a 50% increase in wages for a type #2 person, and a 61% decrease in wages for a type #3 person. The magnitude of these effects indicates that variation in preferences across people is key determinant of occupational choices. The importance of preferences in explaining occupational choices is considered in greater detail in the following chapter.

One way of assessing the relative importance of permanent heterogeneity, match values, and random shocks in determining career choices is to compute the fraction of the unexplained variation in wages and non-pecuniary utility flows attributed to each of the error terms. The results of this decomposition are presented in Table 9. The top half of Table 9 shows the percentage of the unexplained variation in wages in occupation q due to permanent heterogeneity in skills, μ^q, job matching, ψ_{ij}, and random wage shocks, e_{ijt}. The total unexplained variation in wages is simply the sum of the three error components in the model, $\mu_i^q + \psi_{ij} + e_{ijt}$.

The results of this decomposition indicate that occupation-specific skill endowments and job matching are both important determinants of wages. For example, the first row of Table 9 shows that 17% of the unexplained variation in wages in the professional and managerial occupation is due to variation in skill endowments (μ^q) across people, and 24% is due to job matching (ψ_{ij}). The fraction of the unexplained variation in wages attributed to job matching is fairly stable across the five occupations, ranging from 17% to 29%. These results indicate that job matching is a significant determinant of wages in all five occupations. In contrast, the importance of occupation-specific skills varies substantially across occupations. For example, the fraction of unexplained variation in wages attributed to skill heterogeneity ranges from a high of 40% for sales and clerical workers to a low of only 2% for service workers.

The bottom half of Table 9 decomposes the variance of the error term in the non-pecuniary utility flow equation ($\phi_i^q + \xi_{ij} + \varepsilon_{ijt}$) into the fraction due to permanent heterogeneity in preferences, ϕ_{ij}, non-pecuniary job matching, ξ_{ij}, and random utility shocks, ε_{ijt}. The results of this decomposition indicate that permanent heterogeneity in preferences is an even more important determinant of variation in utility flows than permanent heterogeneity in skills. Heterogeneity in preferences for attending school accounts for 65% of the total unexplained variation in the schooling utility flow equation.

The results also show that variation in preferences across people is most important in determining utility for professional and managerial workers and craftsmen, and is least important for sales and clerical workers. The firm-specific non-pecuniary match value accounts for approximately zero percent of the variation in non-pecuniary utility flows across all five occupations. This result suggests that when workers are searching for jobs, the relevant considerations are only their skills in the occupation, the quality of the wage match value, and their preference for working in the occupation. Once these factors are taken into account, the estimates indicate that workers do not attach any importance to firm-specific non-pecuniary job characteristics. In other words, preferences for non-pecuniary job characteristics appear to be occupation-specific, rather than firm-specific.

7 Specification Tests

This chapter examines how well the structural model fits the data. First, simulated educational and employment choices from the structural model are compared to the educational and employment choices observed in the data. A comparison of the simulated and actual choices shows how well the model fits key outcomes in the data. This analysis is extended to include χ^2 goodness-of-fit tests to formally test the null hypothesis that the simulated choice proportions from the model are the same as the choice proportions observed in the data. Finally, the wage distribution generated by the structural model is compared to the one observed in the NLSY data.

7.1 Comparing Simulated and Actual Choices

The simulated choice sequences for the entire sample are used to compute the frequencies of occupational and educational choices for each age. There are $1,023$ simulated choice sequences spanning 15 years, for a total of $15,345$ person-years of simulated data

Figure 1 provides a comparison of the simulated and actual proportion of aggregate years spent employed in each occupation, attending school, and unemployed. In general, the model does a fairly good job of fitting the occupational choice data at this level of aggregation. The proportion of years spent employed as professionals and managers, craftsmen, and service workers are virtually identical in the simulated and actual data. However, the model over predicts the percentage of person-years spent employed as operatives and laborers and sales and clerical workers by 10 and 6 percentage points.

The structural model understates the fraction of person-years spent attending school by 5 percentage points. In the NLSY data, 25% of all years are spent attending school, while only 20% of the years of simulated data are spent attending school. The school attendance proportions translate into an average of 3.15 years spent in school for the NLSY sample, compared to an average of 2.65 years in school for people in the simulated data. The school attendance rate found in the simulated data are primarily influenced by the structural parameters that determine the consumption value of attending school and the pecuniary and non-pecuniary returns to investment in education. The current version of the model allows for three types of people, who have different preferences for attending school. It seems likely that three types of people may not be a flexible enough specification for schooling ability. The current specification of the model allows for one type of person with an extremely low preference for attending school, and two types of people with very high preferences for attending school. Adding a fourth type of person to the model would allow for the possibility of an intermediate level of schooling ability, and may improve the model's ability to fit the aggregate school attendance rate.

The model also under predicts the rate of unemployment. Only 4% of the simulated person-years of data are spent unemployed, while 14% of the years of data in the NLSY are spent in unemployment. Again, this appears to be a situation where allowing for more types of people could improve the fit of the model. The three types of people in the current version of the model all place a relatively high value on working relative to unemployment. These high preferences for employment, combined with the monetary returns to working make unemployment an unattractive option for most people. Adding a fourth type of person to the model who might place a lower value on working and also have lower skills could increase the predicted unemployment rate.

Comparisons of aggregate choice frequencies are a useful first step in assessing how well the model fits the data. However, comparing the data at this high level of aggregation does not provide any information about how well the model matches the trends in people's career choices over time. For example, it is well known that the proportion of individuals attending school declines with age. It would be possible for the model simulations to match the aggregate school attendance rate while missing the strong downward age pattern in school attendance. Figures $2-8$ provide a graphical comparison of the simulated and actual choice frequencies over the sample period. Overall, the model seems to capture the qualitative age patterns in the data quite well, with a few exceptions. The largest discrepancies occur for employment as professionals and managers, employment as operatives and laborers, and unemployment. The model fails to capture the upward age trend in professional employment even though it is able to predict the aggregate proportion of years spent in the professional occupation quite closely, with 22% of simulated years spent as professionals compared to 20% in the actual data. The model closely tracks the age pattern in employment as craftsmen, but overstates employment in the operatives and laborers occupation by approximately 10 percentage points over almost the entire age range. The model over predicts sales and clerical employment for all ages, but tracks the general qualitative age trend.

The current specification of the model allows age to have a direct effect on wages and non-pecuniary utility flows. The wage equations include a linear age term as well as dummy variables for the high school and college age ranges. The specification of the non-pecuniary utility flow equations are more restrictive, because they include only a linear age term. Allowing age to enter the non-pecuniary utility flow equations in a more flexible way may improve the fit of the model to the age pattern in occupational choices. The current version of the model also restricts the effects of observable characteristics such as age and education on employment utility to be the same across occupations. This restriction constrains the effect of age on employment utility to be the same for professional workers and laborers. Relaxing this restriction would improve the ability of the model to match certain elements of the age pattern in occupational choices, such as the increase in professional employment with age.

The model produces a downward age trend in school attendance that is quite similar to the one observed in the data. In general, school attendance declines slightly more smoothly with age in the simulations than in the actual data, and the school attendance rate is always higher in the actual data. The model closely track the unemployment rate observed in the data for 16 and 17 year-olds, but then understates unemployment by approximately 10% in each of the following years.

The preceding analysis has concentrated on examining how well the model fits aggregate choice frequencies and the choice pattern present in the sample over time. Another important feature of the data to consider is mobility between firms and occupations. In the observed data, the average worker moves between firms 2.82 times during the course of a career. In the simulated data, the average number of transitions between firms is 2.08, so the model under predicts mobility between firms. It is difficult to determine exactly why the model would understate the frequency of mobility between firms, but there are several possible explanations. The current specification of the model allows for mobility costs, so changes in this parameter will have a direct impact on mobility between firms. However, the mobility cost is not allowed to vary with observable characteristics. A more flexible specification of the mobility cost function could allow the cost of switching firms to vary with age, for example. This type of change in the mobility cost function would probably improve the ability of the model to match the rate of mobility between firms observed in the data. Another possible explanation for the under prediction of mobility is the fact that the distribution of firm-specific match values is restricted to only three discrete values in the current version of the model. Increasing the number of discrete points would probably increase mobility rates because workers would have more opportunities to move to higher match values.

The patterns of occupational mobility found in the simulated and actual data are compared in Table 10. The top entries in each cell are computed using the actual data, and the bottom entries are computed using the simulated data. The diagonal elements of Table 10 show how well the model captures the persistence in occupational choices found in the NLSY data. The model closely matches the level of persistence in choices for professional and managerial workers, craftsmen, and service workers. In contrast, the model overstates persistence in occupational choices for operatives and laborers (69% actual vs. 87% simulated) and sales and clerical workers (62% actual vs. 83% simulated). The current version of the model does not allow for any utility costs (or benefits) of moving between occupations, so there aren't any parameters that will directly allow the model to match occupational mobility rates in the data. The current version of the model explains occupational mobility in several ways. Workers may choose to move to a new occupation if they have the opportunity to obtain a higher firm match value. However, moving across occupations also destroys occupation-specific human capital. Mobility across occupations is greatly influenced by the correlation of skills and preferences across occupations. For example, people will tend to move from occupations where their wages are low to ones where their wages are higher. In contrast, workers will tend to move in both directions between occupations where their wages are similar. The current version of the model incorporates unobserved heterogeneity in skills and preferences by allowing for three discrete types of people. Using more types of people would allow for a more complete representation of variation in skills and preferences across people, and might allow the model to better fit the patterns of occupational mobility found in the data.

The failure of the model to capture mobility out of the sales and clerical occupation is due in large part to the fact that the model under predicts mobility from sales and clerical jobs into managerial positions by 12 percentage points. The understatement of mobility out of the operatives and laborers occupation is largely due to the fact that the model under predicts mobility in both directions between operatives and craftsmen. One possible explanation for this feature of the simulated data are that the structural model does not allow for transferability of occupation-specific human capital across occupations. If human capital is in fact transferable across occupations, then the model will understate mobility between occupations where human capital has value in both occupations. For example, in the structural model, human capital acquired as an operative has no value in the craftsmen occupation. If human capital acquired as an operative actually has value in the craftsmen occupation, then mobility will be under predicted by the structural model because the model assigns incorrectly low wages to occupation switchers.

7.2 Formal Tests of the Fit of the Model

These tests indicate whether the outcomes predicted by the model are statistically different from those observed in the data. The basic framework applies to a situation where there are K mutually exclusive alternatives. Define P_k as the proportion of observations in the data who choose alternative k, and let $\widehat{P}_k$ represent the model's prediction of that proportion. Let N_k represent the number of observations in each group. The predicted proportions are computed by simulating choice sequences for each person in the sample, and then computing the frequency of each of the K outcomes of interest. The details of the process of simulating choices from the model are discussed in greater detail in the previous chapter. The null hypothesis for this statistical test is that the model is generating choices that are the same as those observed in the actual data. The test statistic is

$$\sum_{k=1}^{K} \frac{N_k(P_k - \widehat{P}_k)^2}{\widehat{P}_k} \sim \chi^2_{K-1,\alpha},$$

where α is the significance level of the test. The null hypothesis is rejected if the value of the test statistic is greater than $\chi^2_{K-1,\alpha}$.

Table 12 presents Chi squared goodness-of-fit test statistics for the aggregate proportions found employed in each occupation, attending school, and unemployed. With one degree of freedom, the critical value from the χ^2 distribution at the .05 level of significance is 3.841, so the null hypothesis that the simulated and actual proportions are equal is rejected if the test statistic is greater than 3.841. The test statistics in Table 12 show that the null hypothesis is rejected for all outcomes except for the employment proportions for craftsmen and service workers. It appears that the model is generating choice proportions for employment in the other three occupations, school attendance, and unemployment that are statistically different from those observed in the data.

Although the model generates outcomes that are statistically different from those observed in the data, it is worth noting that the model does seem to capture many of the important qualitative features of the data, as shown in the previous section. The inability of the model to match certain features of the data may be partly a result of the assumption that heterogeneity in skills and preferences is completely captured by three types of people. The types of people in this model vary across eleven dimensions of ability and preferences, so there is no reason to think that heterogeneity will be sufficiently captured by only three discrete types. In fact, similar discrete choice dynamic programming models of employment and educational choices such as Keane and Wolpin (1997), Eckstein and Wolpin (1999), Stinebrickner (2001), and Keane and Wolpin (2001) use four types of people to model unobserved heterogeneity. In future efforts the number of discrete types of people will be increased from the current level of three. Allowing for more discrete types will definitely improve the fit of the model, but the magnitude of the improvement is uncertain.

7.3 Simulated and Actual Wages

The preceding sections in this chapter have considered how well the model mimics the choice sequences observed in the NLSY data. This section considers how well the model fits the wage distribution observed in the NLSY data. Table 12 shows that the means and standard deviations of accepted wage in the NLSY and simulated data are quite similar. The discrepancy between mean simulated and actual wages ranges from .02 for operatives and laborers to .16 for service workers.

The simulated and actual standard deviations are also quite similar across the five occupations, so it appears that the model generates a wage distribution that is similar to the one observed in the NLSY data.

The model seems to match the mean and variance of the wage distribution fairly well. The following regression of wages on age allows for a comparison of the age pattern in the NLSY data and the simulated data. Standard errors are in parentheses below the coefficients.

$$\begin{aligned}
\text{Actual Data} \quad &: \quad \ln(w) = \underset{(.012)}{9.119} + \underset{(.001)}{.047}Age \\
\text{Simulated Data} \quad &: \quad \ln(w) = \underset{(.012)}{9.073} + \underset{(.001)}{.056}Age
\end{aligned}$$

The regression coefficients show that wages rise about 1% faster with age in the simulated data compared to the NLSY data. The model seems to mimic the general age pattern of accepted wages in the NLSY sample.

8 Counterfactual Experiments

One of the advantages of estimating a structural model relative to reduced form approaches is that the parameters of the structural model can be used to perform counterfactual experiments. If the model is correctly specified, then the parameter estimates reflect the true underlying preferences of agents and the constraints present in the environment in which they are making decisions. These parameter estimates are invariant to policy changes, so they can be used to simulate the behavioral responses of agents to hypothetical changes in the economic environment. Changes in the economic environment are implemented by changing a subset of the structural parameters. The impact of the change in the parameter values is determined by using the new parameter values to simulate career choices. The predicted choice proportions from the maximum likelihood parameter estimates used in the previous chapter to examine the fit of the model now serve as a baseline for comparison with the simulated data from the counterfactual experiment.

One important fact to keep in mind when evaluating the counterfactual simulations is that they are partial equilibrium simulations because they ignore equilibrium considerations. For example, changes in preferences for employment across occupations may result in changes in the wage distribution as the supply of workers in each occupation changes. Similarly, changes in educational policy that alter the supply of workers in different skill categories may alter the equilibrium wage distribution. These equilibrium effects are not considered in the present analysis, since modeling equilibrium is far beyond the scope of this paper.

The first counterfactual experiment is intended to quantify the impact of preferences on occupational choices. In this experiment, heterogeneity in preferences is eliminated from the model by setting the non-pecuniary employment intercepts equal to their mean value, which is calculated across both occupations and types of people. When this restriction is imposed all people have the same mean preferences for employment in each of the five occupations. In this setting, where heterogeneity in preferences is eliminated, occupational choices are driven entirely by wages and non-pecuniary utility shocks. Comparing occupational choices from this experiment with those from the baseline simulation illuminates the role that preferences play in determining occupational choices over the course of a person's career.

The simulated choice frequencies from the baseline and counterfactual simulations are presented together in Figure 9. Figure 9 shows the impact of the elimination of variation in preferences across people on the aggregate choice frequencies. The most significant changes are seen in employment for professionals and managers and operatives and laborers. The proportion of person-years spent in the professional and managerial occupation increases sharply from .22 to .36, while the proportion of years spent as operatives and laborers declines by 12 percentage points. The simulations suggest that preferences cause people to choose to work in other occupations over working as a professional or manager. In the absence of this variation in preferences more people choose to work in the highest paying professional and managerial occupation. Preferences also seem to be an important factor that leads people to choose to work in the relatively low wage operatives and laborers occupation.

The second counterfactual experiment explores the impact of schooling ability (or preference for attending school) on career choices, and examines the extent to which the government can influence occupational choices by subsidizing school attendance. For the purposes of this experiment, suppose that the government is concerned about inequality in schooling ability across people in the population, and is also concerned about how variation in schooling ability influences occupational choices. Suppose that the government decides to provide a subsidy to individuals in each year that they attend school in order to increase educational attainment, and possibly influence occupational choices. In this policy experiment, the government equalizes the mean consumption value of attending school for all people by providing a subsidy of $3,714 for each year of school attendance for type #1 people, $0 for type #2's, and $29 for type #3 people.

In this simulation, all people are assigned a schooling ability parameter (ϕ^s) equal to the highest schooling ability type as a result of the school subsidy policy experiment. That is, the schooling utility flow intercept parameter is increased from 5.19 to 13.41 for type #1 people, and increased from 10.03 to 13.41 for type # 3 people. The impact of the second counterfactual simulation on educational and occupational choices is shown in Figure 10. The largest changes are seen in the proportion of years spent attending school, which nearly doubles. The average number of years that a person spends in school increases from 2.65 to 3.9. Not surprisingly, a schooling subsidy program that increases the consumption value of attending school causes large increases in educational attainment. However, the changes in educational attainment have virtually no impact on occupational choices. Given that the labor market returns to education vary across the five occupations, the lack of interaction between educational attainment and occupational choices is somewhat surprising. For example, one might expect employment in occupations where education is highly rewarded to increase when average educational attainment increases. The fact that this does not happen suggests that variation in schooling ability only plays a minor role in determining occupational choices, although it certainly plays a large role in determining educational attainment. It appears that occupation-specific abilities and preferences are a far more important determinant of occupational choices than schooling ability. This policy experiment suggests that a schooling subsidy program could have a large effect on educational attainment, but is unlikely to alter occupational choices by very much.

9 Conclusion

This paper formulates and structurally estimates a dynamic model of educational, occupational, and inter-firm mobility choices using data from the NLSY. The model developed in this work extends the basic dynamic human capital framework introduced by Keane and Wolpin (1997) to include less highly aggregated occupational groups, firm-specific human capital, and firm-specific wage and non-wage matching. The incorporation of job search over firms within the occupational choice framework allows this work to shed light on the importance of job search over firms relative to occupation-specific skills and preferences in determining wages and career choices. This work also provides estimates of the effects of firm and occupation-specific human capital on wages.

The structural estimates suggest that both firm and occupation-specific human capital play important roles in determining wages. The effects of firm and occupation-specific human capital on wages vary substantially across occupations, which suggests that the degree to which human capital is firm or occupation-specific varies across occupations. Differences in occupation-specific abilities across people are also shown to be a key determinant of occupational choices and wages. The estimates also indicate that preferences for the type of work done in each occupation play a large role in determining people's career choices. Counterfactual simulations show that the effect of preferences on occupational choices is large relative to the effect of variation in skills or schooling ability. Overall, the results suggest that educational and occupational choices are shaped by a complex pattern of comparative advantages in skills and preferences.

In addition to human capital accumulation and variation in skills and preferences, this work shows that firm-specific job matching in wages is also an important determinant of career choices. Occupational choice models that fail to take into account the roles of job matching and firm-specific human capital accumulation will be missing two key factors that determine wages. The estimates also suggest that non-wage job matching is not an important determinant of career choices. Once a person's occupation-specific abilities and preferences are taken into account, the empirical evidence indicates that firm-specific non-pecuniary job characteristics contribute little to utility. In other words, job search over firms seems to be driven by wages, not non-wage job characteristics.

The work presented in this paper also contributes to the growing literature in the area of the estimation of detailed and "realistic" dynamic discrete choice models. The model developed in this paper expands on previous work by incorporating educational choices, occupational choices, job search, and endogenous human capital accumulation within the framework of a dynamic human capital model. Incorporating these features within a unified model permits a comprehensive analysis of the career decision process, but also greatly increases the complexity and computational burden of estimation. Estimation of this model, with its fairly general error structure and extremely large state space is made feasible by employing simulation and interpolation methods when solving the dynamic programming problem. The use of parallel processing techniques, which have only recently become widely available, are also instrumental in making estimation feasible.

Appendix A: Simulation of the Likelihood Function

With the exception of the integral over the distributions of firm and occupation-specific human capital, all integrals are simulated using simple frequency simulators. This type of simulator is not practical in the case of the integral over fc and oc because the distributions of these unobserved state variables are intractably complex. The integral that needs to be evaluated is the path probability over the sample period, denoted Γ. The equation for this probability is

$$\Gamma = \int \int \prod_{t=1}^{\widetilde{T}_i} \Pr[O_{it} \mid \Theta, S_{it}, \Phi_i = \Phi_k, oc, fc] dF(oc) dF(fc).$$

Note that the integral is over the joint distribution of fc and oc over the entire $\widetilde{T}_i$ years that person i remains in the sample. Human capital evolves randomly conditional on career choices, so there are an enormous number of possible sequences of human capital that could occur. Calculating this distribution for each sample person is not practical. The solution is to use a modified GHK algorithm to simulate the integral. The intuition behind this method is the same as in Brien, Lillard, and Stern (2003). The complete algorithm is outlined below.

1. Draw $oc_t^r \mid oc_{t-1}^r$ and $fc_t^r \mid fc_{t-1}^r$.

2. Compute $\Pr[O_{it} \mid oc_t^r, fc_t^r]$.

3. Compute $\Gamma^r = \Gamma^r * \Pr[O_{it} \mid oc_t^r, fc_t^r]$.

4. If $t = \widetilde{T}_i$, go to step 5. Otherwise, set $t = t + 1$ and go to step 1.

5. Repeat these steps for each of the R simulation draws. The simulated path probability is $\Gamma = \frac{1}{R} \sum_{r=1}^{R} \Gamma^r$.

   This algorithm simplifies the problem because drawing fc and oc conditional on the previous draw is very straightforward, while drawing from the complete distribution would be very difficult.

References

[1] Banks, Jeffrey and Rangarajan Sundaram (1994). "Switching Costs and the Gittins Index." *Econometrica*, v. 62: 687-694.

[2] Bartel, Ann (1982). "Wages, Nonwage Job Characteristics and Labor Mobility." *Industrial and Labor Relations Review*, v. 35: 578-589.

[3] Berkovec, James, and Steven Stern (1991). "Job Exit Behavior of Older Men." *Econometrica*, v. 59, no. 1: 189-210.

[4] Berndt, Earnst, Bronwyn Hall, Robert Hall and Jerry Hausman (1974). "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement*, v. 3: 653-665.

[5] Blau, David (1991). "Search for Nonwage Job Characteristics: A Test of the Reservation Wage Hypothesis." *Journal of Labor Economics,* v. 9, Issue 2: 186-205.

[6] Borsch-Supan, Axel and Vassilis Hajivassiliou (1993). "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models." *Journal of Econometrics*, v. 58: 347-368.

[7] Bound, John and Alan Krueger (1991). "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, v. 9: 1-24.

[8] Brien, Michael, Lee Lillard, and Steven Stern (2003). "Cohabitation, Marriage and Divorce in a Model of Match Quality." Unpublished paper, University of Virginia.

[9] Brown, Meta and Christopher Flinn (2002) "Investment in Child Quality Over Marital States." Unpublished Paper.

[10] Cameron, Stephen, and James Heckman (1993). "The Nonequivalence of High School Equivalents." *Journal of Labor Economics*, v. 11, Issue 1, Part 1: 1-47.

[11] Canals-Cerda, Jose, and Steven Stern (1998). "Empirical Search Models." University of Virginia Working Paper.

[12] Christensen, Bent, and Nicholas Kiefer (1994). "Measurement Error in the Prototypal Job-Search Model." *Journal of Labor Economics* v. 12: 618-639.

[13] Eckstein, Zvi, and Kenneth Wolpin (1999). "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." *Econometrica*, v. 67, No. 6.

[14] Ferrall, Christopher (1997). "Empirical Analysis of Occupational Hierarchies." *Journal of Human Resources*, 32: 1-34.

[15] Geweke, John (1988). "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics*, v. 38: 73-89.

[16] Gould, Eric (2002). "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills In the United States." *Journal of Labor Economics*, v. 20, no. 1: pp. 105-147.

[17] Hall, Robert (1982). "The Importance of Lifetime Jobs in the U.S Economy." American Economic Review, v. 72: 716-724.

[18] Heckman, James, and Guilherme Sedlacek (1985). "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self Selection in the Labor Market." *Journal of Political Economy, v. 93: 1077-1125.*

[19] Heckman, James, and Burton Singer (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, v. 52: 271-320.

[20] Jovanovic, Boyan (1979). "Job Matching and the Theory of Turnover." *Journal of Political Economy*, v. 87, no. 5: 972-990.

[21] Kambourov, Gueorgui and Iorii Manovskii (2002). "Occupational Specificity of Human Capital." Unpublished Paper, University of Western Ontario.

[22] Keane, Michael, and Kenneth Wolpin (1994). "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*, v. 76: 648-672.

[23] Keane, Michael, and Kenneth Wolpin (1997). "The Career Decisions of Young Men." *Journal of Political Economy*, v. 105 : 474-521.

[24] Keane, Michael, and Kenneth Wolpin (2001). "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, v. 42, No. 4: 1051-1103.

[25] Light, Audrey (2001) "In-School Work Experience and the Returns to Schooling." Journal of Labor Economics, vol. 19, no. 1: pp. 65-93.

[26] McCall, Brian (1990). "Occupational Matching: A Test of Sorts." *Journal of Political Economy*, v. 98: 45-69.

[27] McCue, Kristin (1996). "Promotions and wage Growth." *Journal of Labor Economics*, v. 14: 175-209.

[28] Miller, Robert A. (1984). "Job Matching and Occupational Choice." *Journal of Political Economy*, v. 92: 1086-1120.

[29] Mincer, Jacob and Boyan Jovanovic (1981), "Labor Mobility and Wages." Studies in Labor Economics, The University of Chicago Press. Edited by Sherwin Rosen.

[30] Murnane, Richard, John Willett, and John Tyler (2000). "Who Benefits from Obtaining a GED? Evidence from High School and Beyond." *The Review of Economics and Statistics*, v. 82: 23-37.

[31] Neal, Derek (1995). "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics, v. 13: 653-677.*

[32] Neal, Derek (1999). "The Complexity of Job Mobility Among Young Men." *Journal of Labor Economics*, v. 17, No. 2.

[33] Parent, Daniel (2000). "Industry-Specific Capital and the Wage Profile: Evidence from the NLSY and PSID." Journal of Labor Economics, v. 18: 306-323.

[34] Roy, Andrew (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, v. 3: 135-146.

[35] Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Howard Zurcher." *Econometrica*, v. 55: 999-1033.

[36] Rust, John (1994). "Structural Estimation of Markov Decision Processes." Handbook of Econometrics, vol. 4, Edited by Robert F. Engle and Daniel L. McFadden. Amsterdam: North-Holland.

[37] Rust, John and Christopher Phelan (1997). "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." Econometrica, v. 65: 781-831.

[38] Sicherman, Nachum (1990). "Education and Occupational Mobility." *Economics of Education Review, v. 9: 163-179.*

[39] Stinebrickner, Todd (2001). "Compensation Policies and Teacher Decisions." *International Economic Review*, v. 42, No. 3: 751-779.

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

[41] Tyler, John, Richard Murnane, and John Willett (2000). "Estimating the Labor Market Signalling Value of the GED." *Quarterly Journal of Economics*: 431-467.

[42] Willis, Robert (1986). "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." Handbook of Labor Economics, v. 1. Amsterdam: North-Holland.

[43] Willis, Robert, and Sherwin Rosen (1979). "Education and Self Selection." *Journal of Political Economy*, v. 100:S7-S36.

[44] Wolpin, Kenneth (1992). "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth." Journal of Political Economy, v. 100: 535-560.

Table 1
Description of Aggregated Occupations

Occupation Group	*Description*	*1970 Census Occupation Codes*	*Example Occupations*
Occupation 1	Professional, Technical, Managers	001 - 245	Architects, Economists, Office Managers
Occupation 2	Craftsmen	401 - 580	Carpenters, Electricians, Automobile Mechanics
Occupation 3	Operatives & Non-farm Laborers	601 - 785	Butchers, Truck Drivers, Groundskeepers
Occupation 4	Sales & Clerical	260 - 395	Insurance Agents, Bank Tellers
Occupation 5	Service & Private Household	901 - 984	Janitors, Dishwashers, Nursing Aides

Table 2
Summary of Occupational Mobility by Age

Ages	*Conditional on Switching Firms, % Switching Occupations*	*Conditional on not Switching Firms, % Switching Occupations*
16-21	56.52%	30.21%
22-25	47.36%	24.58%
26-30	40.70%	15.84%
31-35	37.09%	7.40%
All Ages	47.86%	19.35%

Note: Probabilities are computed using all consecutive years of employment observed in the data for each age group.

Table 3
Occupational Transition Matrix: All Ages

	Occ 1: Professional & Managers	*Occ 2: Craftsmen*	*Occ 3: Operatives & Laborers*	*Occ 4: Sales & Clerical*	*Occ 5: Service*	*Total*
Occ 1: Professional & Managers	83.28	4.22	3.00	7.35	2.15	100.00
Occ 2: Craftsmen	7.25	75.59	13.05	2.55	1.57	100.00
Occ 3: Operatives & Laborers	4.74	14.90	68.98	7.66	3.71	100.00
Occ 4: Sales & Clerical	20.45	4.60	10.76	61.94	2.25	100.00
Occ 5: Service	10.53	7.22	9.32	4.51	68.42	100.00
Total	32.09	22.69	22.43	14.08	8.70	100.00

Note: The entries in this table are transition probabilities from the occupation in the left column to the occupation in the top row. Only consecutive years of employment are used.

Table 4
Occupational Transition Matrix: Ages 16-22

	Occ 1: Professional & Managers	*Occ 2: Craftsmen*	*Occ 3: Operatives & Laborers*	*Occ 4: Sales & Clerical*	*Occ 5: Service*	*Total*
Occ 1: Professional & Managers	66.51	7.08	8.49	13.68	4.25	100.00
Occ 2: Craftsmen	4.62	67.93	20.11	4.89	2.45	100.00
Occ 3: Operatives & Laborers	3.85	17.89	60.54	13.04	4.68	100.00
Occ 4: Sales & Clerical	17.39	7.02	18.73	53.18	3.68	100.00
Occ 5: Service	9.96	9.96	13.28	8.12	58.67	100.00
Total	14.87	24.03	31.24	17.51	12.36	100.00

Table 5
Occupational Transition Matrix: Ages 30-35

	Occ 1: Professional & Managers	*Occ 2: Craftsmen*	*Occ 3: Operatives & Laborers*	*Occ 4: Sales & Clerical*	*Occ 5: Service*	*Total*
Occ 1: Professional & Managers	93.64	2.73	0.91	1.82	0.91	100.00
Occ 2: Craftsmen	6.47	85.29	5.59	1.47	1.18	100.00
Occ 3: Operatives & Laborers	2.70	6.42	86.82	1.69	2.36	100.00
Occ 4: Sales & Clerical	15.48	4.17	4.17	75.60	0.60	100.00
Occ 5: Service	7.14	0.89	4.46	0.00	87.50	100.00
Total	43.27	21.26	18.65	9.45	7.36	100.00

Table 6
Choice Distribution by Age

Age	*Employed*	*School*	*Unemployed*	*Employed and School*	*Employed and GED*	*Unemployed and GED*	*Total*
16	36 3.52	661 64.61	94 9.19	216 21.11	4 0.39	12 1.17	1,023 100.0
17	76 7.89	511 53.06	109 11.32	254 26.38	1 0.10	12 1.25	963 100.00
18	269 30.12	318 35.61	186 20.83	113 12.65	2 0.22	5 0.56	893 100.00
19	347 41.41	251 29.95	167 19.93	69 8.23	0 0.00	4 0.48	838 100.00
20	374 46.87	196 24.56	154 19.30	70 8.77	1 0.13	3 0.38	798 100.00
21	408 53.97	149 19.71	136 17.99	60 7.94	3 0.40	0 0.00	756 100.00
22	476 66.67	71 9.94	115 16.11	46 6.44	4 0.56	2 0.28	714 100.00
23	502 74.37	32 4.74	100 14.81	39 5.78	2 0.30	0 0.00	675 100.00
24	518 80.81	27 4.21	66 10.30	26 4.06	3 0.47	1 0.16	641 100.00
25	505 83.20	10 1.65	72 11.86	19 3.13	0 0.00	1 0.16	607 100.00
26	504 85.57	11 1.87	50 8.49	23 3.90	0 0.00	1 0.17	589 100.00
27	483 85.94	5 0.89	59 10.50	14 2.49	1 0.18	0 0.00	562 100.00
28	451 84.14	11 2.05	56 10.45	16 2.99	1 0.19	1 0.19	536 100.00
29	440 85.27	2 0.39	67 12.98	4 0.78	1 0.19	2 0.39	516 100.00
30	446 89.56	3 0.60	47 9.44	2 0.40	0 0.00	0 0.00	498 100.00
Total	5,835 55.55	2,258 21.28	1,478 13.93	971 9.15	23 .21	44 .41	10,609 100.00

Table 7
Structural Model Estimates

Variable	*Occupations* *Professional & Managers*	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*
Log Wage Equation:					
Age (β_1)	-.039* (.001)	-.082* (.002)	-.064* (.0005)	-.046* (.0005)	-.049* (.0009)
Years of High School (β_2)	-.043* (.004)	-.108* (.004)	.003* (.0008)	.272* (.001)	-.033* (.004)
Years of College (β_3)	.193* (.002)	.058* (.002)	.078* (.0005)	.355* (.0008)	.072* (.003)
Age $\leq$ 17 (β_4)	-.704 (.072)	-1.290* (.056)	-.804* (.033)	-.018* (.051)	-.718* (.048)
18 $\leq$ Age $\leq$ 21 (β_5)	-.404* (.029)	-.620* (.023)	-.539* (.019)	-.370* (.026)	-.475* (.029)
GED (β_6)	.228* (.025)	-.034* (.016)	.100* (.003)	.743* (.005)	-.497* (.023)
Firm-specific HK: Level # 1 (β_7)	0.00&	0.00&	0.00&	0.00&	0.00&
Firm-specific HK: Level # 2 (β_8)	.090 (.072)	.3295* (.050)	.1541* (.033)	.6256* (.004)	.276* (.070)
Firm-specific HK: Level # 3 (β_9)	.722* (.028)	.3295* (.026)	.1541* (.031)	.6256* (.026)	.432* (.034)
Occupation-specific HK: Level # 1 (β_{10})	0.00&	0.00&	0.00&	0.00&	0.00&
Occupation-specific HK: Level # 2 (β_{11})	.009 (.201)	.001 (.003)	.125* (.014)	0.00 (0.001)	0.00 (.001)
Occupation-specific HK: Level # 3 (β_{12})	.820* (.023)	.330* (.017)	.7485* (.006)	1.07* (.016)	2.66* (.029)
Probability that Human Capital Increases (λ)	.15&	.15&	.15&	.15&	.15&

Notes: *Denotes statistically significant at the 5% level.
& Indicates parameters that are fixed at the stated value. These parameters are not estimated.
Age is measured as true age minus 15.

Table 7
Structural Model Estimates, Continued

Variable	*Estimate*	*Standard Error*
Discount Factor (δ)	.95[&]	
School Utility Flow		
Age (γ_{s1})	-2.356*	.036
Attending College Dummy (γ_{s2})	1.893*	.184
Attending Graduate School Dummy (γ_{s3})	-4.373*	.189
Years of High School (γ_{s4})	-.423*	.050
Years of College (γ_{s5})	.469*	.043
School While Employed Utility Flow		
Age (γ_{sw1})	-1.539*	.034
Years of High School (γ_{sw4})	-.542*	.051
Years of College (γ_{sw5})	.559*	.065
GED Utility Flow		
Constant (γ_{g1})	-13.549*	.611
Age (γ_{g2})	-.581*	.072
Non-Wage Employment Utility		
Age (α_1)	.785*	.006
Years of High School (α_2)	-.295*	.003
Years of College (α_3)	-.507*	.004
High School Diploma (α_4)	1.079*	.024
College Diploma (α_5)	1.141*	.015
GED (α_6)	.314*	.005
Occupation-specific HK (α_8)	.893*	.004
Firm-specific HK (α_7)	1.397*	.024
Switching Costs		
Cost of Moving to a New Firm (firm to firm transitions) (α_{10})	3.769*	.064
School Re-entry Cost (γ_{s6})	-.869*	.129
Cost of Moving to a new Job from Non-Employment (α_9)	2.775*	.111
Costs of Working while Attending School		
Work in High School	3.539*	.194
Work in College	3.987*	.209
Work in Graduate School	5.647*	.274

Notes: * Denotes statistically significant at the 5% level.

Table 7
Structural Model Estimates, Continued

Variable	*Estimate*	*Standard Error*
Within-firm Job Offer Probabilities		
Offer from Occupation # 1: Professional & Managers (π_1)	.36*	.013
Offer from Occupation # 2: Craftsmen (π_2)	.24*	.015
Offer from Occupation # 3: Operatives & Laborers (π_3)	.17*	.009
Offer from Occupation # 4: Sales & Clerical (π_4)	.14*	.007
Offer from Occupation # 5: Service (π_5)	.08*	.004
Error Standard Deviations		
True Randomness in Wages (σ_e)	.355*	.002
Non-Pecuniary Firm Match Value (σ_ξ)	.016	.095
Pecuniary Firm Match Value (σ_ψ)	.227*	.012
Measurement Error in Wages (σ_ζ)	0.000	0.00
Extreme Value Parameter (τ)	2.224*	.009

Notes: *Denotes statistically significant at the 5% level.

Table 7
Structural Model Estimates, Continued

Variable	***Type #1***	***Type #2***	***Type #3***
Log-wage Intercepts			
Professional & Managerial (μ^1)	9.963 [1] (.027)	9.551 [3] (.025)	9.944 [2] (.035)
Craftsmen (μ^2)	10.460 [3] (.023)	10.607 [2] (.034)	10.862 [1] (.031)
Operatives & Laborers (μ^3)	9.982 [2] (.014)	9.722 [3] (.017)	10.183 [1] (.025)
Sales & Clerical (μ^4)	8.662 [1] (.019)	7.914 [3] (.012)	8.369 [2] (.058)
Service (μ^5)	9.748 [3] (.028)	9.859 [2] (.029)	9.890 [1] (.045)
Non-pecuniary Intercepts			
Professional & Managerial (ϕ^1)	-11.673 [3] (.063)	-10.583 [2] (.046)	-4.959 [1] (.037)
Craftsmen (ϕ^2)	-10.026 [2] (.034)	-11.276 [3] (.057)	-4.015 [1] (.058)
Operatives & Laborers (ϕ^3)	-9.280 [3] (.015)	-8.607 [2] (.017)	-4.251 [1] (.028)
Sales & Clerical (ϕ^4)	-9.952 [3] (.035)	-8.751 [2] (.012)	-6.805 [1] (.154)
Service (ϕ^5)	-10.624 [3] (.205)	-10.594 [2] (.040)	-4.951 [1] (.050)
School (ϕ^s)	5.191 [3] (.025)	13.419 [1] (.276)	10.33 [2] (.402)
Type Probabilities	.59 (.029)	.34 (.027)	.07 (.013)
Log-likelihood	-25,642		

Note: numbers in brackets represent the ranking of each type. Standard errors are in parentheses.

Table 8
Returns to Occupation & Firm Specific Human Capital: Means & Standard Deviations

	2 years	*6 years*	*10 years*
Professional & Managerial			
Firm	.039 (.111)	.197 (.284)	.360 (.332)
Occupation	.021 (.121)	.189 (.340)	.377 (.405)
Craftsmen			
Firm	.091 (.147)	.205 (.160)	.265 (.131)
Occupation	.007 (.049)	.074 (.137)	.151 (.164)
Operatives & Laborers			
Firm	.043 (.069)	.096 (.075)	.124 (.061)
Occupation	.048 (.119)	.217 (.290)	.384 (.336)
Sales & Clerical			
Firm	.174 (.280)	.390 (.303)	.502 (.248)
Occupation	0.024 (.158)	.239 (.446)	.487 (.533)
Service			
Firm	.080 (.131)	.207 (.171)	.293 (.161)
Occupation	0.060 (.394)	.595 (1.110)	1.21 (1.350)

Notes: The top entry in each cell is the mean return to tenure at a firm or in an occupation. The second entry is the standard deviation of the return to tenure.

Table 9
Decomposition of the Variance in Wages & Non-pecuniary Utility Flows

Log-wage Equation	% of Variance Due to Permanent Heterogeneity in Skills (μ)	% of Variance Due to Wage Match Value (ψ)	% of Variance Due to Random Wage Shock (e)
Professional & Managerial	17%	24%	58%
Craftsmen	7%	27%	66%
Operatives & Laborers	10%	26%	64%
Sales & Clerical	40%	17%	42%
Service	2%	29%	69%
Non-pecuniary Utility	**% of Variance Due to Permanent Heterogeneity in Preferences (ϕ)**	**% of Variance Due to Non-pecuniary Match Value (ξ)**	**% of Variance Due to Random Utility Shock (ε)**
Professional & Managerial	26%	0%	74%
Craftsmen	27%	0%	73%
Operatives & Laborers	16%	0%	84%
Sales & Clerical	9%	0%	91%
Service	20%	0%	80%
School	65%	•	35%

Table 10: Actual and Simulated Occupational Transition Matrix: Actual Data (top entry) & Simulated Data (bottom entry)

	Professional & Managers	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*	*Total*
Professional & Managers	83.28 84.57	4.22 4.15	3.00 4.63	7.35 4.84	2.15 1.81	100.00
Craftsmen	7.25 7.71	75.59 78.72	13.05 8.35	2.55 3.61	1.57 1.61	100.00
Operatives & Laborers	4.74 4.01	14.90 4.69	68.98 87.21	7.66 2.62	3.71 1.48	100.00
Sales & Clerical	20.45 8.36	4.60 3.52	10.76 3.84	61.94 83.24	2.25 1.05	100.00
Service	10.53 10.28	7.22 5.31	9.32 6.7	4.51 5.2	68.42 72.52	100.00
Total	32.09 28.96	22.69 17.48	22.43 29.85	14.08 16.87	8.70 6.84	100.00

Note: The entries in this table are transition probabilities from the occupation in the left column to the occupation in the top row. Only consecutive years of employment are used. The top entry in each cell is computed from the actual data, and the bottom entry is from the simulated data.

Table 11: Goodness-of-Fit Specification Tests

	Actual Proportion in the NLSY	***Simulated Proportion***	***Actual - Simulated***	***χ^2 Test Statistic***
Professional & Managers	.20	.22	-.02	31.1
Craftsmen	.15	.15	0	1.2
Operatives & Laborers	.17	.27	-.10	1063.3
Sales & Clerical	.10	.16	-.06	508.2
Service	.07	.07	0	3.5
Attending School	.25	.20	.05	169.0
Unemployed	.14	.04	.10	1175.3

Notes: Predicted proportions are based on 15,345 simulated person-years, obtained from 1,023 simulated people over 15 years.

Table 12: Wage Distribution: Actual & Simulated Data

	NLSY Data	*Simulated Data*
Professional & Managerial		
Mean	9.77	9.87
Standard Deviation	.535	.637
Craftsmen		
Mean	9.58	9.47
Standard Deviation	.453	.509
Operatives & Laborers		
Mean	9.37	9.35
Standard Deviation	.453	.462
Sales & Clerical		
Mean	9.51	9.62
Standard Deviation	.507	.785
Service		
Mean	9.25	9.41
Standard Deviation	.473	.504

Notes: Predicted proportions are based on 15,345 simulated person-years, obtained from 1,023 simulated people over 15 years.

Figure 1

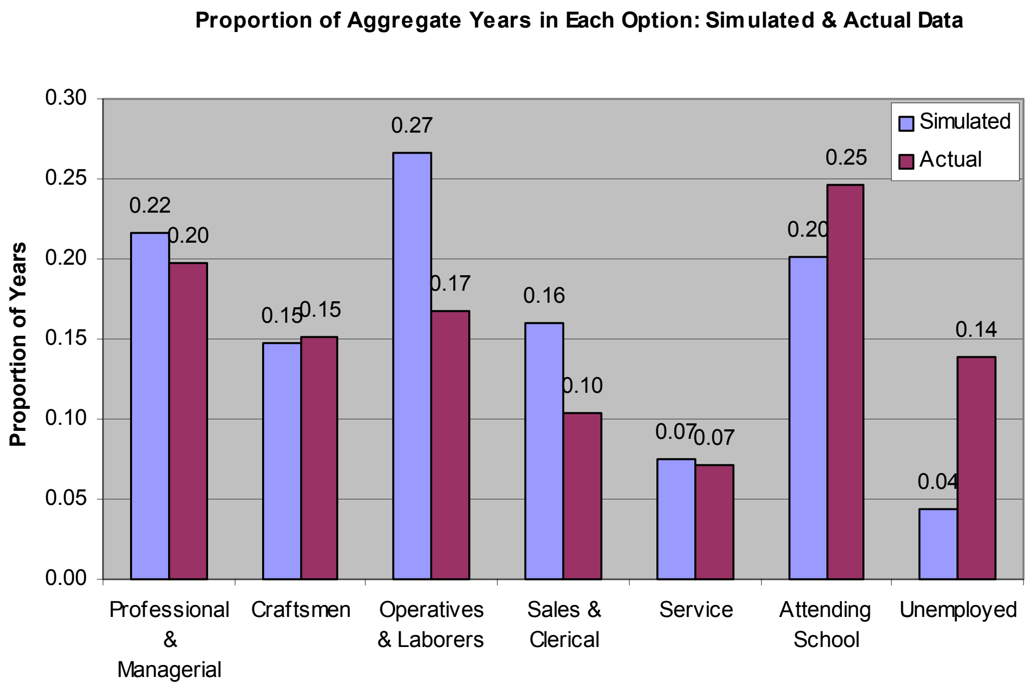

Figure 2

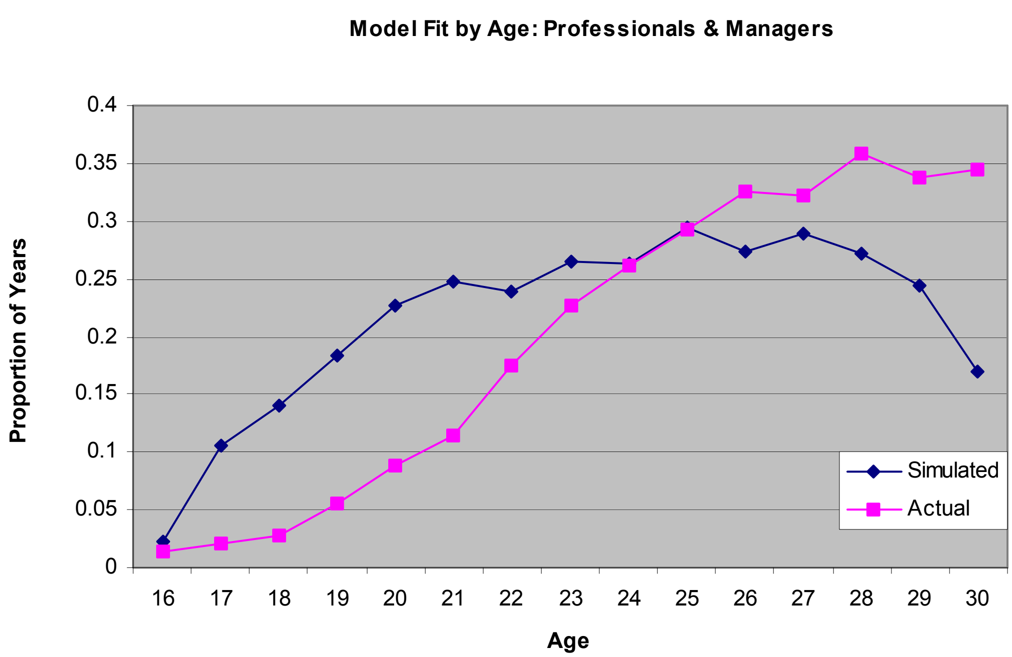

Figure 3

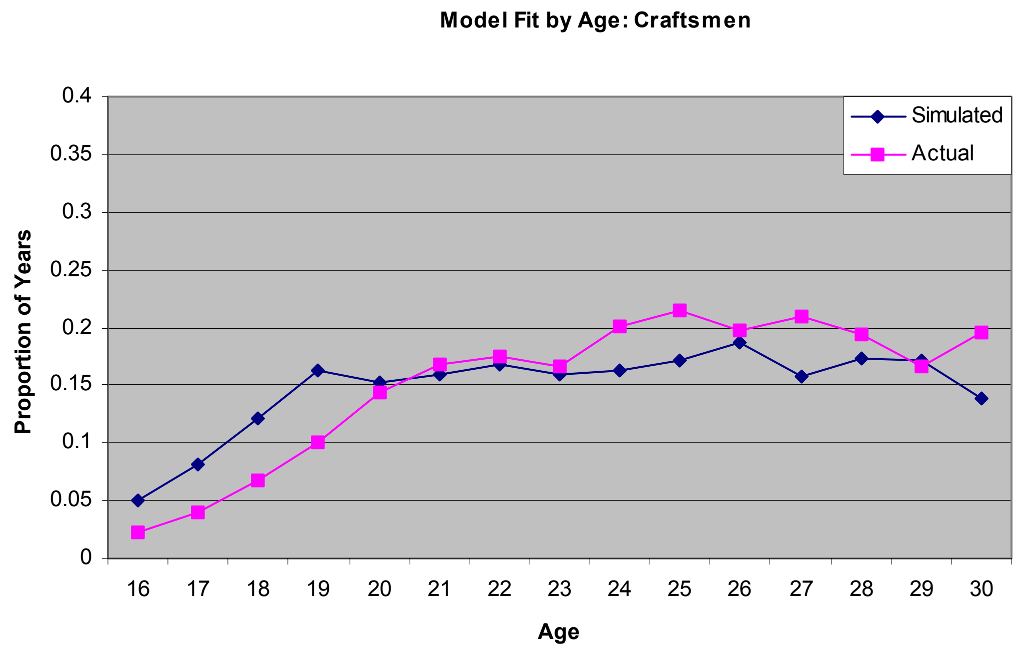

Figure 4

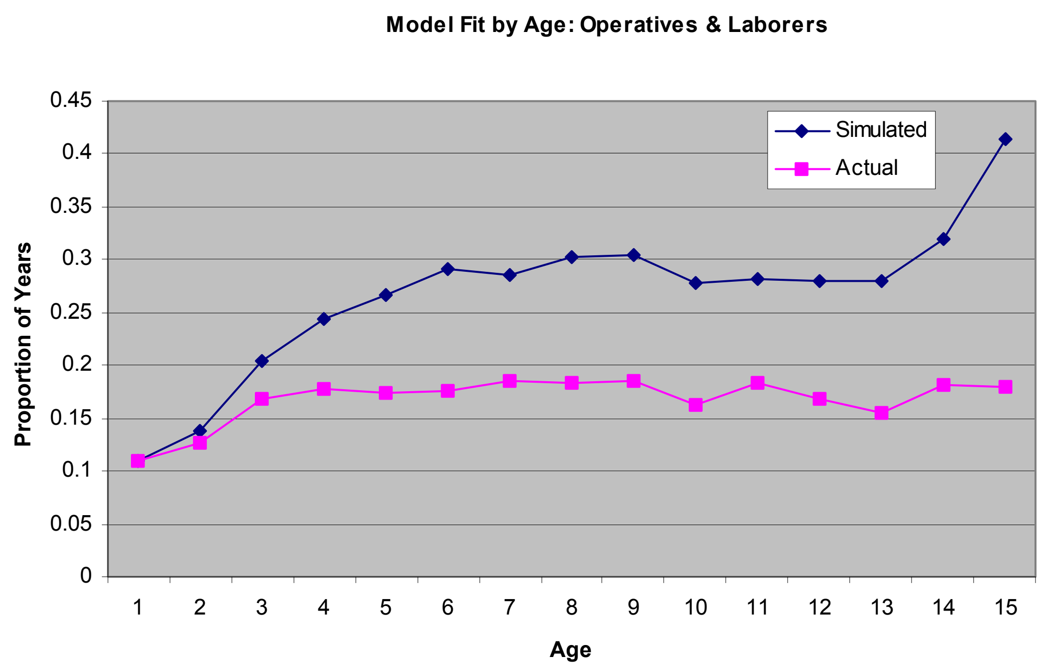

Figure 5

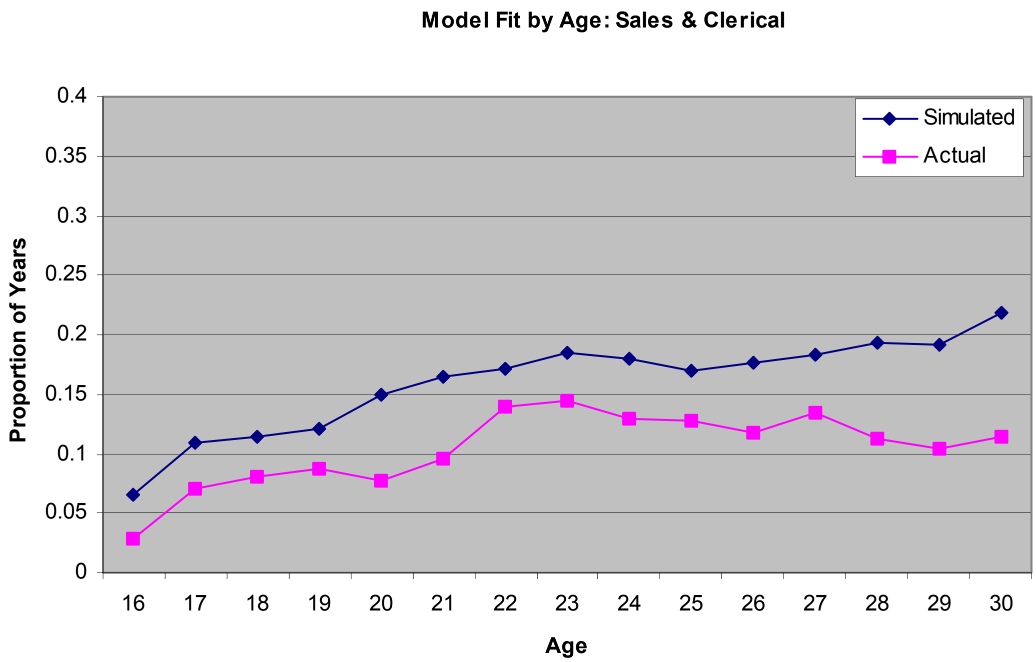

Figure 6

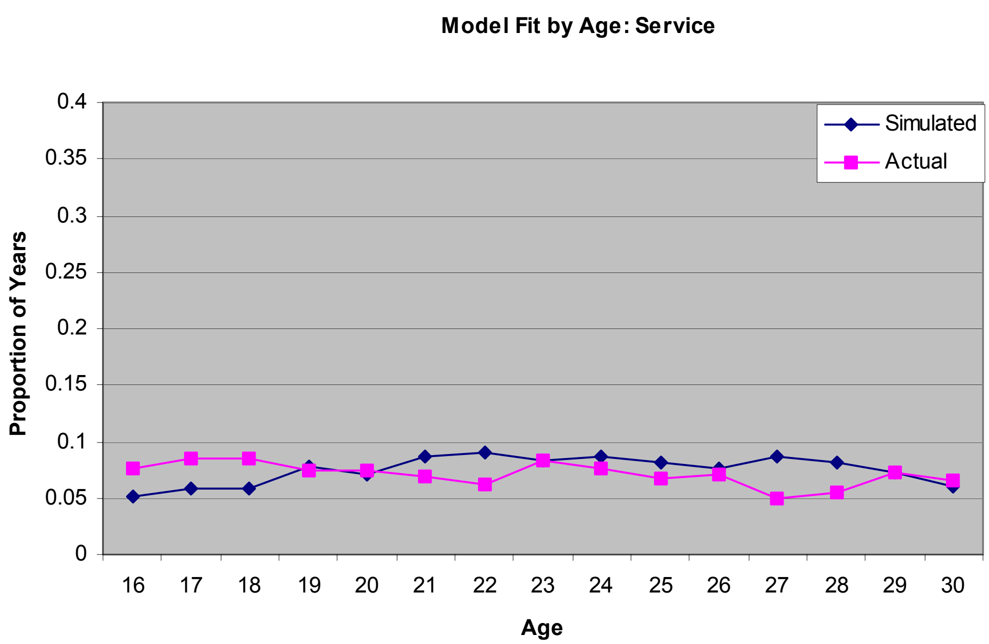

Figure 7

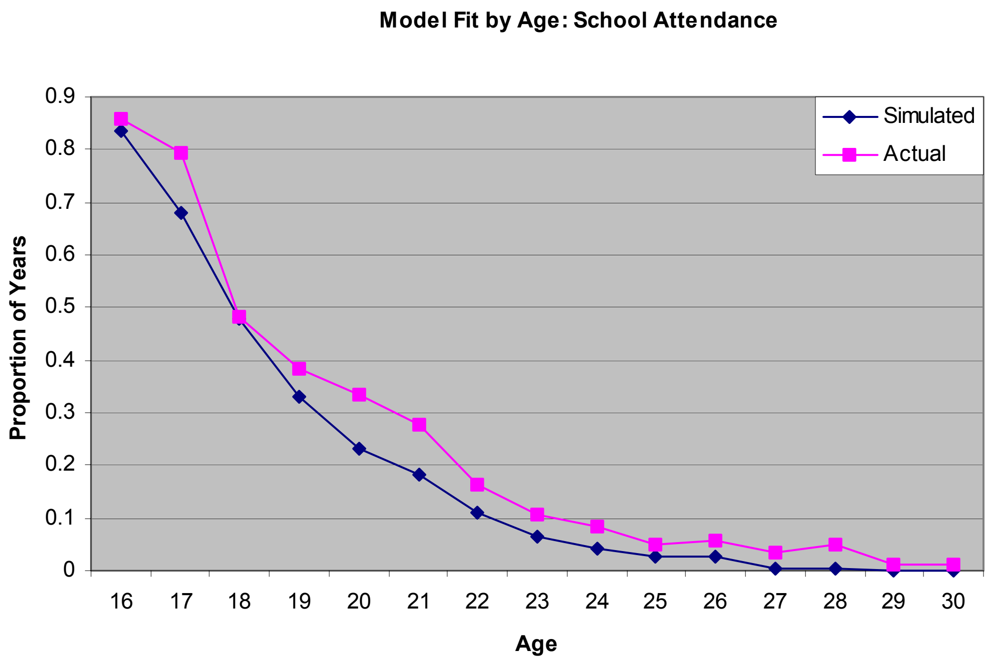

Figure 8

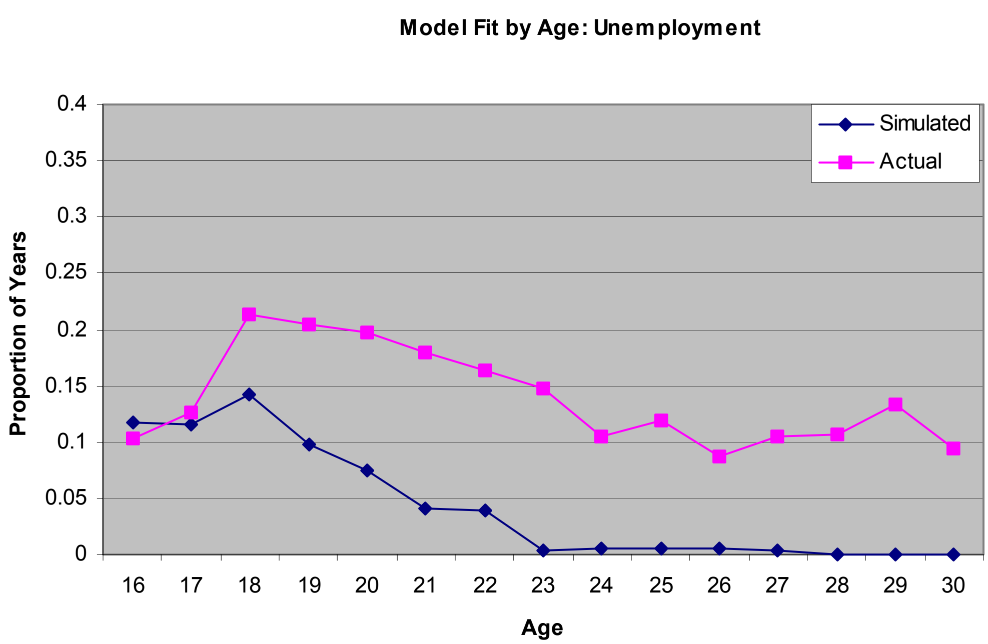

Figure 9

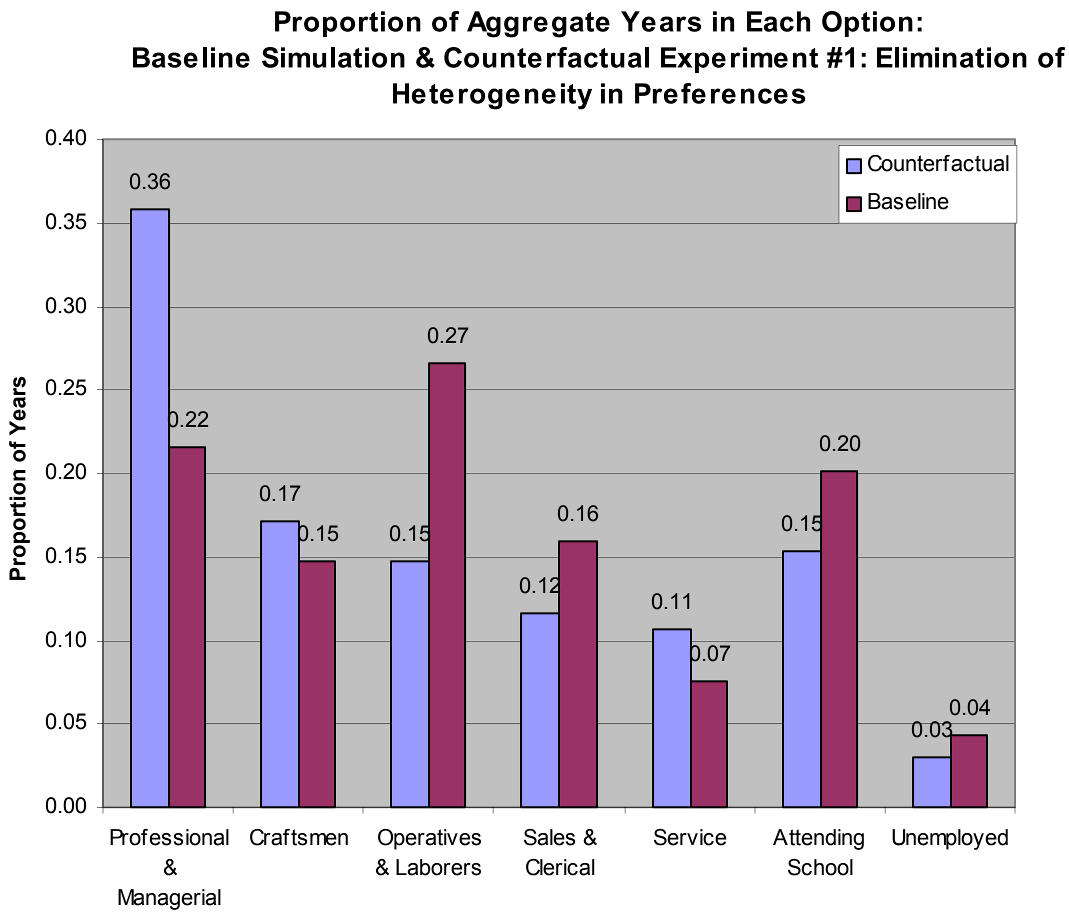

Figure 10

Proportion of Aggregate Years in Each Option:
Baseline & Counterfactual Experiment #2 - Increase in Schooling Ability