A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search

Sullivan, Paul

Bureau of Labor Statistics

September 2006

A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search

Paul Sullivan*
Bureau of Labor Statistics

September 2006

Abstract

This paper examines career choices using a dynamic structural model that nests a job search model within a human capital model of occupational and educational choices. Individuals in the model decide when to attend school and when to move between firms and occupations over the course of their career. Workers search for suitable wage and non-pecuniary match values at firms across occupations given their heterogeneous skill endowments and preferences for employment in each occupation. Over the course of their careers workers endogenously accumulate firm and occupation specific human capital that affects wages differently across occupations. The parameters of the model are estimated with simulated maximum likelihood using data from the 1979 cohort of the National Longitudinal Survey of Youth. The structural parameter estimates reveal that both self-selection in occupational choices and mobility between firms account for a much larger share of total earnings and utility than the combined effects of firm and occupation specific human capital. Eliminating the gains from matching between workers and occupations would reduce total wages by 30%, eliminating the gains from job search would reduce wages by 19%, and eliminating the effects of firm and occupation specific human capital on wages would reduce wages by only 2.7%.

Keywords: Occupational Choice, Job Search, Dynamic Discrete Choice, Human Capital

---

*I would like to thank Steven Stern, John Pepper, and William Johnson for their invaluable assistance during this project. I would also like to thank Loren Smith, John Bound, and numerous seminar participants for helpful comments and suggestions. I gratefully acknowledge financial support for this research provided by the Bankard Fund for Political Economy at the University of Virginia, and the University of Virginia College of Arts and Sciences. Contact information: Email: Sullivan.Paul.Joseph@bls.gov. Phone: (202) 691-6593. All views and opinions expressed in this paper are those of the author and do not necessarily represent the views of the Bureau of Labor Statistics.

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 firms. Despite the interrelated nature of these choices, previous research has generally examined educational and occupational choices separately from decisions about job search. Empirical studies of occupational and educational choices are frequently based on the framework of human capital models, which have taken the form of dynamic programming models in recent work (Keane and Wolpin 1997, Lee 2005, Lee and Wolpin 2006). In these dynamic human capital models workers endogenously accumulate education and occupation specific human capital as they make optimal career choices, but all jobs are identical within an occupation. In contrast to dynamic human capital models, an extensive job search literature has emphasized the importance of job matching between workers and firms in determining wages while generally abstracting away from both occupational choices and human capital accumulation.[1] McCall (1990) and Neal (1999) depart from this trend by developing search models that incorporate occupations, but these models do not include human capital accumulation.

The goal of this research is to further the understanding of how people make decisions about educational attainment and employment by estimating a dynamic structural model of career choices that incorporates the key features of a job search model within a dynamic human capital model of occupational and educational choices. In the model, forward looking workers choose when to attend school and when to move between occupations and firms as they maximize their discounted expected utility. Search frictions such as randomness in job offers and moving costs impede the optimal allocation of workers across occupations and firms. Over the course of their careers workers endogenously accumulate general human capital in the form of education as well as occupation and firm specific human capital. The value of employment varies over the five occupations in the

[1] Berkovec and Stern (1991) and Wolpin (1992) develop search models that include firm specific capital but these models do not incorporate occupational choices.

economy because workers have heterogeneous skill endowments and preferences for employment across occupations, and because the effect of human capital on wages varies across occupations. Workers search for suitable wage and non-pecuniary match values at firms across occupations given their innate skills and preferences and stock of human capital. Allowing for search based on non-pecuniary utility generalizes the approach used in many search models which assume that workers search only for wage match values.[2]

This paper contributes to a growing literature that demonstrates the value of using dynamic discrete choice models to study employment and educational choices over the career. Most relevant to the work presented here are the dynamic human capital models developed by Keane and Wolpin (1997, 2001) and dynamic structural models such as Eckstein and Wolpin (1999) and Belzil and Hansen (2002) that focus on the endogenous accumulation of education. The model developed in this paper expands on the occupational choice model of Keane and Wolpin (1997) by adding job matching between workers and firms, firm specific human capital, heterogeneity in preferences for employment in each occupation, and by expanding the number of civilian occupations from two to five. Incorporating the human capital occupational choice approach to career dynamics along with the firm based job search approach within a unified model is necessary to determine the relative importance of each modelling approach in explaining career choices, wages, and total utility.

The parameters of the structural model are estimated by simulated maximum likelihood using data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY). The likelihood function follows directly from the recursive numerical solution to each individual's dynamic programming problem. The computational cost of estimation is substantial because simultaneously modelling human capital accumulation, job search, and occupational choices creates a dynamic programming problem that is challenging to solve. Estimation is made feasible by implementing an interpolation method when solving the dynamic programming problem that modifies the approach

[2] See Blau (1991), Hwang, Mortensen, and Reed (1998), and Dey and Flinn (2005) for examples of search models that incorporate non-pecuniary job characteristics.

developed by Keane and Wolpin (1994) in a way that takes advantage of the special structure of this model.

The career choice model nests a human capital model of occupational choices and a job search model, so the structural parameter estimates provide direct evidence about the relative importance of firm and occupation specific human capital, job search, and heterogeneity in occupation specific skills and preferences in determining career choices, wages, and total utility. The parameter estimates show that features of a dynamic human capital model and job search model are both necessary to understand the evolution of wages and utility over the career. The potential percent increase in wages caused by the acquisition of firm specific human capital ranges from 9.5% to 24.8% across occupations, and the potential wage gains from occupation specific human capital range from essentially zero to 17% across occupations.[3] The estimates of the effects of human capital on wages indicate that search models that do not incorporate human capital accumulation are missing an important source of wage growth. The potential wage gains from job search are also quite large. A worker who is able to move from a 25th percentile firm match to a 75th percentile match realizes a 45% increase in wages. Occupational choice models set in a human capital framework where all jobs are identical within occupations are missing a key determinant of wage growth.

The structural model is used to conduct counterfactual simulations that quantify the importance of firm and occupation specific capital, education, matching between workers and firms, and matching between workers and occupations in determining total earnings and overall welfare. Previous research has not been able to determine the relative importance of these factors because existing models have not incorporated job search, occupational matching, and human capital accumulation within a unified model. The counterfactual simulations reveal that eliminating the returns to firm and occupation specific capital reduces total wage earnings by 2.7%. The wage gains resulting from

[3] Estimating the return to firm tenure has been the subject of a large literature. See, for example, Altonji and Shakotko (1987), Topel (1991), and Dustmann and Meghir (2005). Estimating the returns to occupation tenure has received far less attention. See, for example, Kambourov and Manovskii (2006) for instrumental variables estimates of the returns to occupation tenure.

utility maximizing mobility between firms and occupations far exceed the wage gains from human capital accumulation. The counterfactual simulations demonstrate that eliminating the wage gains from job search reduces earnings in the simulated economy by 19%, while eliminating the gains from matching between workers and occupations reduces total earnings by 30%. The total social value of matching between workers and firms, which includes pecuniary and non-pecuniary utility, is 33% of total utility. The importance of both job search and occupational choices in determining total earnings and utility highlights the value of jointly modelling these two key facets of the career decision process.

The model is also used to determine the importance of permanent heterogeneity in skills and preferences in determining lifetime utility relative to the importance of randomness in wage and utility shocks, randomness in human capital improvement, and randomness in the arrival of firm-specific job matches. The results indicate that 64% of the variation in lifetime utility is determined by permanent heterogeneity. To provide some context for this result, Keane and Wolpin (1997) find that 90% of lifetime utility is determined by permanent heterogeneity in their career choice model that does not consider the role of matching between workers and firms in determining wages and utility. Allowing for matching between workers and firms, firm specific human capital, and random shocks to non-pecuniary utility reduces the relative importance of permanent heterogeneity in determining lifetime utility, but its impact is still substantial.

The remainder of the paper is organized as follows. The next section discusses the data. Section 3 presents the model of career choices, and Section 4 discusses estimation. Section 5 presents the structural parameter estimates, and Section 6 presents the results of the counterfactual experiments. Section 7 concludes.

2 Data

The parameters of the model are estimated using the 1979 cohort of the National Longitudinal Survey of Youth (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. 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 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.[4] 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

[4] Yearly data is frequently used when estimating dynamic structural models. See, for example, Keane and Wolpin (1997) or Belzil and Hansen (2002).

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.[5] 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

[5] Incorporating part time employment as a choice variable in the structural model is conceptually straightforward, but it would increase the cost of estimation substantially.

aggregation is to compare the average number of jobs that a person holds over the sample period 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.[6]

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, so using extremely detailed occupational classifications is not computationally feasible. Based on these considerations, occupations are aggregated into the five occupational groups listed in Table 1. Aggregating occupations into five groups is a lower level of aggregation than that found in existing research. Recent dynamic structural models of occupational choices such as Keane and Wolpin (1997) and Lee (2005) have aggregated the data into only two occupations (blue and white collar). Lee and Wolpin (2006) model both sectoral and occupational choices by allowing workers to choose between blue, white, and pink collar employment in both the service and goods sectors, but they do not model job search decisions.

[6] 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.

2.1 Descriptive Statistics

This section highlights the key characteristics of the data and provides descriptive statistics about the career choices observed in the data. Table 2 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 86% of the sample attends school at age 16. School attendance takes a discrete drop to 48% at age 18, the age where most people have graduated from high school. As an alternative to high school graduation, 6.6% of the sample reports earning a GED at some point over the sample period. School attendance declines steadily throughout the college ages and then drops to approximately 16% at age 22, the normal college graduation age. School attendance declines to 4.8% by age 25, and continues to decline at more advanced ages. Keane and Wolpin (2001) report a qualitatively similar relationship between age and schooling using less highly aggregated data that divides each school year into three segments. As school attendance declines with age, the percentage of people employed as professional and managerial workers steadily increases from 1.4% at age 16 to 34.5% at age 30. In contrast, the percentage of people employed as service workers is relatively stable over time, ranging between 5 and 8%.

The percentage of people unemployed is 10% at age 16. Unemployment rises to approximately 20% at ages 18-21 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.

Table 1
Description of Aggregated Occupations

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

Table 2
Choice Distribution by Age

Age	*School*	*Professional & Managers*	*Craftsmen*	*Operatives & laborers*	*Sales & clerical*	*Service*	*Unemployed*	**Total Observations**
16	85.7	1.4	2.2	10.9	2.9	7.6	10.4	1,023
17	79.4	2.1	4.0	12.7	7.1	8.5	12.6	963
18	48.3	2.8	6.8	16.9	8.0	8.5	21.4	893
19	38.2	5.6	10.1	17.7	8.8	7.4	20.4	838
20	33.3	8.9	14.3	17.4	7.8	7.4	19.7	798
21	27.6	11.5	16.8	17.6	9.5	6.9	18.0	756
22	16.4	17.5	17.5	18.6	13.9	6.2	16.4	714
23	10.5	22.7	16.6	18.4	14.4	8.4	14.8	675
24	8.3	26.1	20.1	18.6	12.9	7.6	10.5	641
25	4.8	29.2	21.4	16.3	12.7	6.8	12.0	607
26	5.8	32.6	19.7	18.3	11.7	7.1	8.7	589
27	3.4	32.2	21.0	16.9	13.5	5.0	10.5	562
28	5.0	35.8	19.4	15.5	11.2	5.4	10.6	536
29	1.2	33.7	16.7	18.2	10.5	7.2	13.4	516
30	1.0	34.5	19.5	17.9	11.4	6.6	9.4	498
All	24.6	19.8	15.1	16.8	10.4	7.1	13.9	10,609

Note: Entries are percentages. Rows need not sum to 100% because school attendance and employment are not mutually exclusive.

Table 3 shows that there are differences in the levels of inter-firm and intra-firm occupational mobility. The relevant entries in each cell for this discussion are the top entries, which are computed using the NLSY data.[7] 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 29% from the youngest age group to the oldest, while intra-firm occupational mobility declines by 41%. 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.

Table 4 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. The relevant entries for this discussion are the top entries, which are computed using the NLSY data. For example, cell (2,1) indicates that a person employed as a craftsman has a 7.25% chance of becoming a professional or managerial worker in the next year, conditional on being employed in the next year. The diagonal elements of the occupational transition matrix in Table 4 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 as professional and managerial workers are least likely to switch occupations, while sales and clerical workers are most likely to switch occupations.

Overall, the transition matrix is fairly symmetric, with the exception of the flows of workers between the sales and clerical and service occupations to professional and managerial employment. Workers are much more likely to switch from sales and clerical or service employment to professional and managerial jobs than in the opposite direction. The largest flow of workers between occupations occurs from sales and clerical to professional and managerial employment.

[7] The bottom entries in the cells in Tables 3 and 4 are computed using simulated data generated from the estimated structural model. These entries will be discussed in detail later in the paper.

Table 3
Summary of Occupational Mobility by Age: NLSY Data (top entry) and Simulated Data (bottom entry)

Ages	*Conditional on Switching Firms, % Switching Occupations*	*Conditional on not Switching Firms, % Switching Occupations*
16-21	57.64% 54.36%	29.94% 27.42%
22-25	50.09% 47.21%	26.85% 23.51&
26-30	40.76% 37.84%	17.61% 14.98%
All Ages	49.78% 46.51%	24.69% 21.82%

Note: Probabilities are computed using all consecutive years of employment observed in the data for each age group. The top entry in each cell is computed using the NLSY data, and the bottom entry is computed using simulated data generated using the estimated structural model.

Table 4
Occupational Transition Matrix: NLSY Data (top entry) and Simulated Data (bottom entry)

	Professional & Managers	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*	**Total**
Professional & Managers	83.28 86.21	4.22 2.80	3.00 2.51	7.35 6.48	2.15 2.00	100.00
Craftsmen	7.25 5.50	75.59 77.41	13.05 12.17	2.55 4.30	1.57 .62	100.00
Operatives & Laborers	4.74 4.73	14.90 13.56	68.98 71.10	7.66 7.51	3.71 3.10	100.00
Sales & Clerical	20.45 17.01	4.60 5.98	10.76 8.84	61.94 65.73	2.25 2.44	100.00
Service	10.53 8.78	7.22 7.02	9.32 8.10	4.51 6.24	68.42 69.86	100.00
Total	32.09 29.70	22.69 21.35	22.43 22.99	14.08 16.12	8.70 9.84	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. The top entry in each cell is computed using the NLSY data, and the bottom entry is computed using simulated data generated using the estimated structural model.

3 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. Workers search for suitable wage and non-wage match values across firms while employed and non-employed given their skills and preferences for employment in each occupation. Dual activities such as simultaneously working and attending school are also 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 both firm and occupation specific work experience and education, which affect 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.

[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 2 suggests that that a significant fraction of workers switch occupations without switching firms.

3.1 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, $d_t(s) = 1$ indicates that schooling is chosen at time t, and $d_t(s, e3, nf) = 1$ indicates attending school (s) 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 3.1.2.

3.1.1 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, $w_{it}(k)$, and non-pecuniary utility, $H_{it}(k)$, that person i receives from choice combination k at time t,

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

[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 exam on the first try. Within two years the eventual pass rate is 85%.

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

3.1.1a *Wages.* The log-wage of worker i employed at firm j in occupation q at time t is

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

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}. This term 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}) are observed by the worker when a job offer is received.[12]

3.1.1b *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

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

[11] Keane and Wolpin (1997) show that heterogeneity in skill endowments is 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.

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. This term 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

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

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 permanent attributes of employment at each firm that affect the employment utility flow and are not observed by the econometrician. 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, $C^g(S_{it})$.

3.1.2 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.[15]

[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 search 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, since this addition would render an already complicated model completely intractable. Keane and Wolpin (2001) present evidence that although borrowing constrains are severe, relaxing these constraints has little impact on educational attainment. Cameron and Taber (2003) examine this issue using a number of different empirical approaches and consistently find that given the current policy regime, there is no evidence of inefficiencies in the schooling market created by borrowing constraints.

[15] In this model workers always have the option of returning to their current job, although the offered wage will change because each job receives a new random shock in each year (e_{ijt}). Thus, transitions into unemployment are utility maximizing responses to shocks. This framework is adopted in many papers such as Berkovec and Stern (1991), Keane and Wolpin (1997), and Lee and Wolpin (2005). An alternative framework allows for a job destruction (layoff) probability and allows workers to always stay at the existing job at the previous wage. Eckstein and van den Berg (2006) discuss these two alternative models and conclude that "There is no aspect of the data that would force us to prefer one of the two models . . . conceptually the two models are observationally equivalent using data on

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

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

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 education.

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.[16] 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

employment, unemployment, and the observed dynamic transitions among these states and the observed accepted wages."

[16] One possible althernative specification of the model would be to incorporate job offer probabilities from each occupation so that in any given time period, an individual may not receive job offers from all five occupations. One could then attempt to estimate these offer arrival rates. However, because only accepted wages and transitions are observed in the data, this alternative model is completely observationally equivalent to the one presented in this paper. For example, suppose that in the data we observe that laborers rarely ever become professionals. When we only observe transitions and accepted wages it is impossible to distinguish a world where laborers infrequently receive professional job offers from a world where laborers frequently receive very low professional wage offers.

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.

Within-firm occupation switch 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 occupation switch 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.[17]

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.

3.1.3 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

[17] See Canals and Stern (2002) for a discussion of a similar identification issue that arises in a simple search model.

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 due to the fact that the model incorporates job search, occupational choices, and educational choices. In order to keep the model tractable, 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.[18] However, age effects are included in the wage equations to proxy for general human capital that has value in more than one 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

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

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.[19] There are separate skill increase probabilities for firm and occupation specific capital, and the rates of skill increase are also allowed

[18] Ideally, one would allow for cross-occupation experience effects in the wage equation, which would require adding measures of previous occupation specific human capital to the state space. Unfortunately, allowing for these effects would render an already extremely computationally demanding estimation problem completely infeasable given current computer technology. Due to the size of the state space, along with the large number of parameters in the model, estimating the model presented in this paper is only possible using interpolation methods and parallel processing techniques. In addition, it is necessary to have access to a large parallel computing cluster to estimate the model in any reasonable amount of time. Given these considerable difficulties, I leave the inclusion of cross-occupation experience effects as an extension for future research.

[19] Brown and Flinn (2004) use a similar method to model the process by which child quality changes over time.

to vary across occupations. The skill increase parameters are $\{\lambda_f^k, \lambda_o^k, \quad k = 1, ..., 5\}$, where the subscripts f and o refer to firm and occupation specific capital, and k indexes occupations. The human capital transition probabilities (λ's) are 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. The size of the state space is significantly reduced when P is a small number relative to the possible values of years of work experience, but the model still captures the human capital improvement process. In this work, $P = 3$. Sections 5.1 and 5.2 present evidence that the discrete approach to modelling human capital provides parameter estimates that fit the observed patterns in wage growth in the NLSY extremely well.[20]

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

[20] In a note about this paper Pavan (2006) comments that this model only allows for "restrictive wage growth patterns." However, in fact this approach to modeling human capital is quite flexible relative to approaches commonly used in the literature, such as a quadratic specification. This approach allows wage growth profiles to be concave or convex, and heterogeneous across occupations. In addition, the model allows for ex post heterogeneity in the returns to human capital because "unlucky" individuals may never transition to a higher level of human capital over their career, but "lucky" individuals may experience rapid wage growth. Random human capital improvements also impact mobility because unlucky individuals will be more likely to switch jobs that lucky individuals. Pavan (2006) also allows for ex post heterogeneity by assuming that firm and occupation specific human capital improvements follow a random walk with drift. However, one key difference between the two approaches is that in Pavan's model at the start of a worker's career there is no ex ante heterogeneity in expected wage growth because expected wage growth is the same across all occupations. In this model there is ex ante heterogeneity because the returns to human capital are allowed to vary across occupations, and workers have different probabilities of working in each occupation because they have heterogeneous abilities and preferences.

allows for the possibility that years of work experience may vary for people with a given level of human capital.

3.2 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 the current values of all variables in the state vector, S_t, as well as the probability that human capital will increase in the next period, conditional on employment (λ_f^k, λ_o^k, $k = 1, ..., 5$).

The maximization problem can be represented in terms of alternative specific value functions which give the lifetime discounted expected 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 and utility flow equations are functions of the state vector, S_t, but this argument is suppressed for brevity of notation.

The value function for an individual with discount factor δ employed in occupation q is the utility flow from employment, plus the discounted 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.[21]

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 choice,

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

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

[21] 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.

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

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

The corresponding expected value of the maximum term is

$$\begin{aligned} EZ_t^{su} &= E\max\{V_{t+1}(s), V_{t+1}(u), V_{t+1}(u,g), \\ &\quad V_{t+1}(ei,nf), V_{t+1}(m,ei,nf),\ m=s,g,\ i=1,...,5\}, \end{aligned} \tag{11}$$

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 11. He then chooses the option with the highest discounted expected value.[22]

[22] A related model is developed by Pavan (2006), who presents a model of search across firms and careers. Pavan must make a number of assumptions to make his model tractable because he defines careers using very detailed three-digit occupation and industry codes. Some of the key differences between the two papers are: 1) His model imposes the restriction that expected wage growth profiles are identical across all occupations. In contrast, the results in this paper suggest that expected wage growth due to human capital accumulation varies widely across occupations, and that this heterogeneity has important implications for wage growth. 2) Education is exogenous in Pavan's model, and also is restricted to have the same effect on wages across all occupations. These assumptions simplify the model, but as a result his model is unable to account for the fact that educational and occupational choices are made jointly. 3) There is no on the job search in Pavan's model, while there is on the job search in the model presented in this paper. In his model workers must choose to leave their jobs before observing a new offer, while in contrast this paper allows workers to observe their current job offer and the outside job offers before making a mobility decision. In Pavan's model workers will leave bad matches, but the model is unable to account for wage growth resulting from workers receiving an extremely good sequence of job matches, because once workers have found a good match they will be unwilling to leave that match to take a chance on receiving a better offer. As a result, his model predicts that 2/3 of all jobs end exogenously. 4) In Pavan's model, at the start of the career a worker has the same expected wage

3.3 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 11. Before considering the solution of the model in more detail, it is useful to specify the distributions of the random components of utility flows.

3.3.1 Distributional Assumptions

Assume that firm specific match values and randomness in wages are distributed i.i.d normal: $\xi_{ij} \backsim N(0, \sigma_\xi^2)$, $\psi_{ij} \backsim N(0, \sigma_\psi^2)$, and $e_{ijt} \backsim N(0, \sigma_e^2)$. 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.

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

in every occupation, because he does not model comparative advantages across occupations. As a result, all sorting across occupations and mobility between occupations in his model is caused by random wage shocks. Therefore, it is unlikely that his model is able to match the distribution of workers across occupations, mobility patterns between occupations, or the wage distribution in each occupation. 5) Pavan's model does not allow for non-pecuniary utility to impact career choices. 6) His model does not incorporate unemployment. Still, Pavan's model is richer than mine in two dimensions. First, he considers the more general topic of career choice, rather than the occupational choices considered in this paper. He also uses the very detailed three-digit occupation and industry codes to define careers, as opposed to five broad occupational groups. It is likely that using five broadly defined occupational categories understates mobility. However, it should be noted that one problem with using three-digit codes is that they are known to be extremely noisy. For example, a validation study by Mellow and Sider (1983) finds that only 58% of three-digit occupation codes are correctly recorded, so using three-digit codes will lead to a large number of false transitions between occupations.

tribution 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.

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 (1984), Keane and Wolpin (1997), and Stinebrickner (2001), the joint distribution of skill endowments and preferences is 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.

3.3.2 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.[23] 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 dynamic models such as Miller (1984) and Berkovec and Stern (1991),

[23] See, for example, Rust (1987) and Rust and Phelan (1997).

and is the basis of job search models.

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 in both wage and non-pecuniary utility 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,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \mid \Omega_t, \Phi_i, oc_t, fc_t\} &= \tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t \mid \Omega_t, \Phi_i, oc_t, fc_t)}{\tau})]) \\ &= \Psi(d_t \,|\Omega_t, \Phi_i, oc_t, fc_t) \,, \end{aligned} \tag{12}$$

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 for each endowment type,

$$E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \mid \Phi_i\} = \int \int \left[\int \cdots \int \Psi(d_t \,|\Omega_t, \Phi_i, oc_t, fc_t) f(\Omega_t) d\Omega_t \right] f(fc_t) dfc_t f(oc_t) doc_t. \tag{13}$$

This integral does not have an analytical solution, so it is simulated using R draws from the joint density $f(\Omega_t)$. In this work, $R = 40$.[24] 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 13 over the R draws,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon | \; \Phi_i\} \quad = \quad & \frac{1}{R} \sum_{r=1}^{R} \sum_{h=1}^{P} \Pr[fc_t = fc_t(h) \,|\, fc_{t-1}] \sum_{z=1}^{P} \Pr[oc_t = oc_t(z) \,|\, oc_{t-1}] \times \\ & \Psi(d_t \,| \Omega_t^r, \Phi_i, oc_t^z, fc_t^h). \end{aligned} \tag{14}$$

The other Emax terms found in the value function calculations are also 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, and Stern (2006) 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 using a regression at the remaining points in the state space. This paper implements a new regression function that takes advantage of the assumption that the error term 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

[24] 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.

shown in equation 12. 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,

$$
\begin{aligned}
E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon\ \} &= \omega_{0t} + \omega_{1t}\tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t)}{\tau})]) \qquad (15) \\
&= \omega_{0t} + \omega_{1t}\Psi(d_t)\ .
\end{aligned}
$$

The parameters ω_{0t} and ω_{1t} are estimated by OLS, and allowed to vary over time. This regression function has the desirable 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 McFadden (1981). Another important property of this regression function is that the regressor is defined at every point in the state space even if the set of feasible state points varies over 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.[25]

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 .95 and .99. Experimentation shows that the actual and interpolated value functions differ by approximately 1% on average.

4 Estimation of The Structural Model

The parameters of the model are estimated by simulated maximum likelihood (SML) 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

[25] One solution to this problem would be to use a different interpolating regression for each feasible choice set in the state space. Depending on the exact details of the model, this approach has two potential drawbacks: 1) small sample sizes in each individual regression, 2) the need to estimate a large number of interpolating regressions.

the methods used to maximize the likelihood function.

4.1 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.

4.1.1 Wage and 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

$$\begin{aligned} w_q(S_{it}) \quad = \quad & \beta_1^q age_{it} + \beta_2^q age_{it}^2/100 + \beta_3^q hs_{it} + \beta_4^q col_{it} + \beta_5^q 1[age_{it} \leq 17] + \\ & \beta_6^q 1[age_{it} \geq 18 \cap age_{it} \leq 21] + \beta_7^q ged_{it} \\ & +\beta_8^q 1[fc_{it} = fc(1)] + \beta_9^q 1[fc_{it} = fc(2)] + \beta_{10}^q 1[fc_{it} = fc(3)] \\ & +\beta_{11}^q 1[oc_{it} = oc(1)] + \beta_{12}^q 1[oc_{it} = oc(2)] + \beta_{13}^q 1[oc_{it} = oc(3)]. \end{aligned} \tag{16}$$

The parameters β_8^q and β_{11}^q are fixed at zero since they are not separately identified from the constant in the wage equation.[26]

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

$$\begin{aligned} \theta_q(S_{it}) \quad = \quad & \alpha_1^q age_{it} + \alpha_2^q age_{it}^2/100 + \alpha_3^q (hs_{it} + col_{it}) + \alpha_4^q oc_{it} + \alpha_5^q fc_{it} + \alpha_6^q hd_{it} \\ & +\alpha_7^q cd_{it} + \alpha_8^q ged_{it} + \alpha_9^q 1[L_{it} > 5] + \alpha_{10}^q NF_{it} \qquad\qquad q = 1, ..., 5. \end{aligned} \tag{17}$$

The inclusion of explanatory variables in the employment non-pecuniary employment flow equations allows observable variables to have a direct impact on employment utility in addition to any

[26] The inclusion of direct age effects in the wage equations follows the approach of Keane and Wolpin (1997). They allow for a linear age effect along with an age≤17 dummy variable.

effect that they may have on wages. For example, as people age it may be the case that physically demanding occupations become less desirable relative to white collar employment.[27] The cost function for attending school is

$$
\begin{aligned}
c^S(S_{it}) &= \gamma_{s1} age_{it} + \gamma_{s2} age_{it}^2/100 + \gamma_{s3} hd_{it} + \gamma_{s4} cd_{it} + \gamma_{s5} hs_{it} + \gamma_{s6} col_{it} + \gamma_{s7} 1[L_{it} \neq 6] \\
c^{SW}(S_{it}) &= \gamma_{sw1} age_{it} + \gamma_{sw2} age_{it}^2/100 + \gamma_{sw3} hs_{it} + \gamma_{sw4} col_{it} + \gamma_{s7} 1[L_{it} \neq 6] \\
&\quad + \gamma_{sw6}(hs_{it} \leq 4) + \gamma_{sw7}(hs_{it} = 4 \cap col_{it} \leq 4) + \gamma_{sw8}(col_{it} \geq 4). \qquad (18)
\end{aligned}
$$

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 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.[28]

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

$$c^g(S_{it}) = \gamma_{g1} + \gamma_{g2} age_{it}. \qquad (19)$$

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

$$\pi_j = \frac{\exp(\rho_j)}{\sum_{k=1}^{5} \exp(\rho_k)}. \qquad (20)$$

[27] See for example, Stinebrickner (2001) for an example of a dynamic discrete choice model that allows for similar effects of observable variables on non-pecuniary utility from working in a teaching job relative to a non-teaching job. Identification of these parameters follows from the use of wage data during estimation. For example, suppose that after controlling for wage differentials across occupations, higher educated people choose a certain occupation more frequently than one would expected based only on wages. This suggests that education impacts the employment non-pecuniary utility flows.

[28] The specification of the schooling utility flow equation is based closely on Keane and Wolpin (1997). One of Keane and Wolpin's (1997) major findings is that a "bare bones" dynamic human capital model that excludes age effects and re-entry costs from the schooling utility flow equation is unable to match the rapid decine in schooling with age. Including direct age effects of this sort has become standard in the dynamic human capital literature. In addition, it seems reasonable to believe that the effort cost of schooling (or non-pecuniary consumption value) varies with age.

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.[29]

4.2 The Likelihood Function

The likelihood function used to estimate the structural model follows directly from the model presented in Section 3. 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,

$$\begin{aligned} L_i(\Theta \mid \Phi_i = \Phi_k) \quad = \quad & \int \cdots \int [\int \int \left(\prod_{t=1}^{\widetilde{T}_i} \Pr[O_{it} \mid \Theta, S_{it}, \Phi_i = \Phi_k) \right) \\ & dF(oc) dF(fc)] dF(\Omega). \end{aligned} \tag{21}$$

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.

The high dimensional integrals in the likelihood function are simulated using R draws from

[29] 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, their estimate is .936.

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 simulated likelihood contribution is

$$L_i^R(\Theta \mid \Phi_i = \Phi_k) = \frac{1}{R}\sum_{r=1}^{R}\frac{1}{Q}\sum_{q=1}^{Q}\prod_{t=1}^{\widetilde{T}_i}\Pr[O_{it}^{rq} \mid \Omega_i^r, oc^q, fc^q, \Theta, S_{it}, \Phi_i = \Phi_k). \tag{22}$$

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

$$L_i^R(\Theta) = \sum_{m=1}^{M}\chi_m L_i^R(\Theta \mid \Phi_i = \Phi_m). \tag{23}$$

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

$$L^R(\Theta) = \prod_{i=1}^{N} L_i^R(\Theta). \tag{24}$$

The vector of parameters $\widehat{\Theta}$ that maximizes equation number 24 is the simulated maximum likelihood estimate of Θ.

4.2.1 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,

$$\Pr(d_{it} = s \,|\, \Omega,\, oc, fc, \Theta, S_{it}, \Phi_i) = \frac{\exp(V_t(s))}{\sum\limits_{k \in D_t^{ne}} \exp(V_t(k))}. \tag{25}$$

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 the distributions of Ω, oc, and fc as discussed previously. The probabilities for outcomes involving employment are similar to the non-employed outcome probabilities, except the choice probability is conditioned on the observed wage and multiplied by the wage density.

The parameters of the structural model are estimated using a derivative based optimization routine. The covariance matrix of $\widehat{\Theta}$ is estimated using the "outer product of the gradient" method of Berndt, Hall, Hall, and Hausman (1974). Estimation of the model challenges the limits of currently available computers due to the extremely large state space of the model, but estimation is made feasible by employing parallel processing techniques.

5 Structural Parameter Estimates

Tables B1-B4 in Appendix B present the structural parameter estimates and the associated standard errors. This section discusses selected parameter estimates and their implications for the career decision process.

5.1 Model Fit

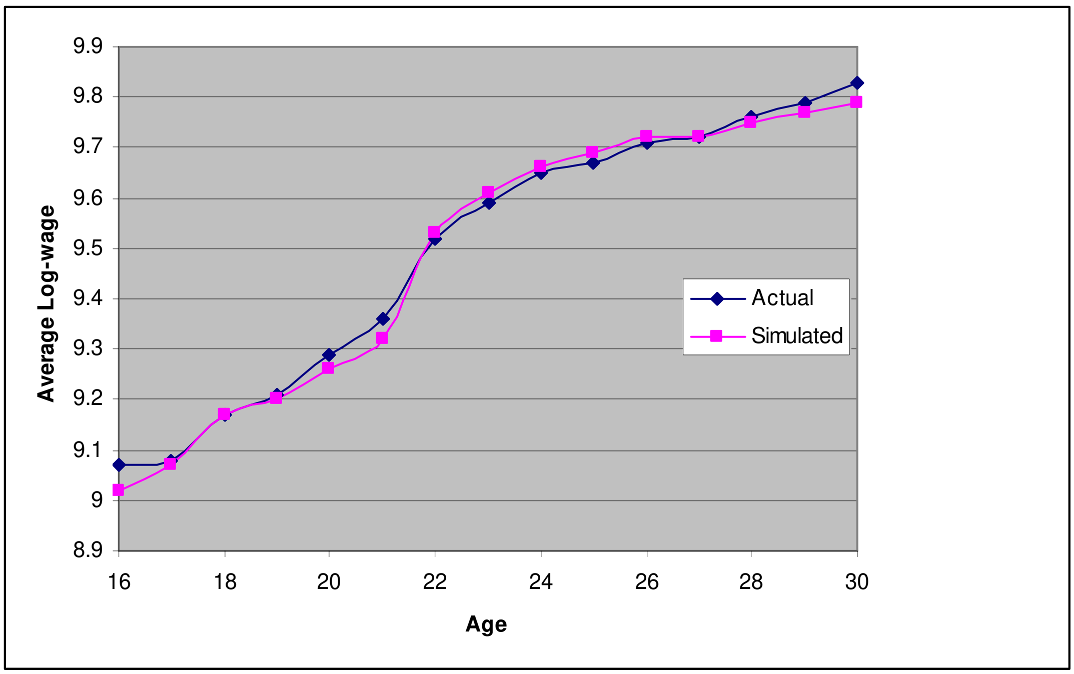
Before discussing the parameter estimates it is useful to consider how well the model is able to match the patterns found in the NLSY career choice and wage data. The structural parameter estimates are used to simulate a sample of $2,000$ individuals whose career choices and wages are compared to those observed in the data. The results of this exercise are presented in Figures 1-2 and Tables 3-5. Table 5 shows the means and standard deviations of accepted log wages in the

Table 5
Wage Distribution: Actual & Simulated Data

Variable	*Professional & Managers*	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*
Mean wage: NLSY data	9.78	9.58	9.37	9.51	9.25
Mean wage: simulated data	9.80	9.61	9.38	9.59	9.33
Wage std dev: NLSY data	.535	.453	.453	.507	.473
Wage std dev: simulated data	.504	.458	.431	.483	.463

Note: Simulated wages computed from a sample of 2,000 people. Yearly wages are in logs.

Figure 1: Actual vs. Simulated Mean Log-wages

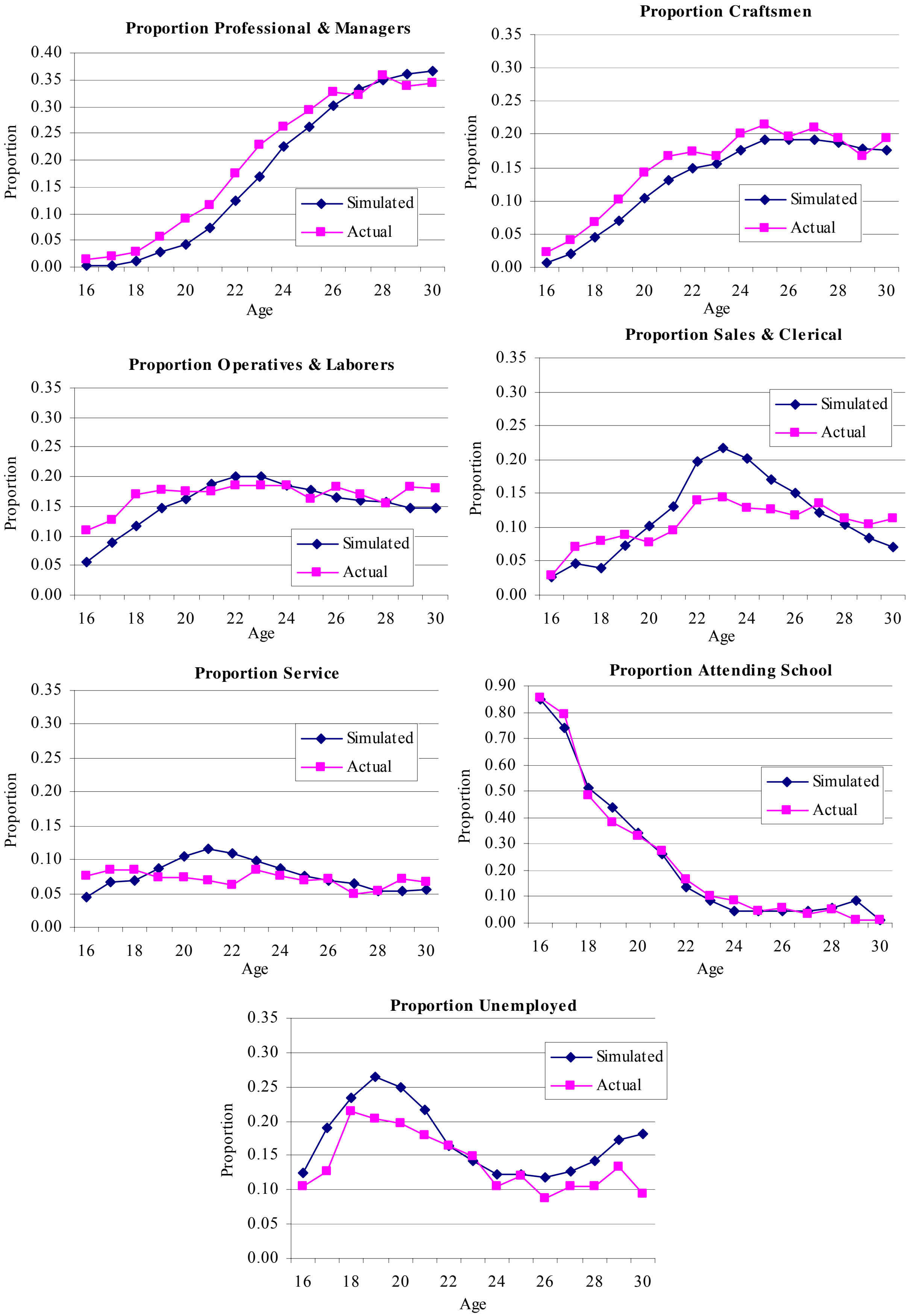


NLSY and simulated data. The discrepancies between simulated and actual mean log wages range from .01 to .08 across occupations, and the model matches the standard deviations of wages in each occupation even more closely. The model fits the observed wage distribution quite well. In addition, Figure 1 shows that the model is able to match the age profile of wages extremely closely. The model captures the general upward trend in mean wages and the sharp increase in mean wages that occurs at college graduation quite precisely.

Tables 3 and 4 show how well the model fits the patterns of occupational mobility found in the NLSY data. In Tables 3 and 4 the top entry in each cell is computed using the NLSY data, and the bottom entry is computed using the simulated data. Table 3 shows that the model is able to match the rates of inter-firm and intra-firm occupational mobility extremely well. The model captures the fact that inter-firm occupational switching is more common that intra-firm occupational switching, and the model also matches the sharper downward age trend in intra-firm occupational mobility. Table 4 shows that the model is also able to closely match the occupational transition matrix found in the NLSY data, so the model generates patterns in occupational mobility that are quite similar to those found in the NLSY. The diagonal elements of Table 4 show that overall, the model tends to slightly overstate persistence in occupational choices, but in general the model's fit to occupational mobility is quite good.

Figure 2 depicts the choice proportions disaggregated by age for both the NLSY data and simulated data. The model fits the choices observed in the data quite well, in most cases closely tracking both the levels of the choice proportions found in the NLSY data as well as the age trends. The model closely matches the sharp upward age trend in professional and managerial employment found in the NLSY data, and the model also matches the more gradual increase in craftsmen employment with age. The model also captures the relatively flat age patterns in the operatives and service occupations. The model tracks the downward age trend in school attendance extremely closely. The simulated data reproduces the general qualitative age pattern in unemployment found

Figure 2: Choice Proportions by Age – Actual vs. Simulated Data



in the NLSY data, although the model over predicts the unemployment rate in the late teens and early twenties. The model also overstates employment in the sales and clerical occupation during the mid twenties. Overall, a comparison of the actual and simulated data shows that a model that combines features of a job search model and human capital occupational choice model is able to fit the wage distribution extremely well, and is also able to closely match patterns in occupational and educational choices.

5.2 The Log Wage Equation: Human Capital & Job Search

The estimates of the log wage equation parameters found in Table B1 reveal the importance of education and occupation and firm specific human capital in determining wages in each occupation. The effects of high school and college on wages vary widely across the five occupations, which suggests that the types of skills produced by high school and college education are valued differently across occupations. The percent change in wages resulting from completing an additional year of high school ranges from a low of 1.3% for craftsmen to a high of 5.9% for operatives and laborers. The monetary return to attending college also varies across occupations. A year of college increases wages by approximately 9.7% for professional, managerial, and service workers, while a year of college increases wages by only 3.1% for an operative or laborer. The effects of education on wages are statistically significant at the 5% level in all occupations. The relationship between education and wages is convex in four out of the five occupations, with only operatives and laborers realizing a lower wage gain from college education than high school education.

The finding that the wage function is convex in schooling differs from the results of most studies of the relationship between schooling and wages which typically assume linearity (Card 1999). A notable exception is Belzil and Hansen (2002) who also find a convex schooling-wage function based on their estimates of a dynamic programming model of schooling and employment choices. In the present model, the average return per year of education from grade ten to sixteen is 7.6% for professional and managerial workers, 3.2% for craftsmen, 4.3% for operatives and laborers, 5.4% for sales and clerical workers, and 6.4% for service workers. These results are consistent with the relatively low average return to schooling of 7% per year reported by Belzil and Hansen (2002), given that they do not allow the returns to schooling to vary by occupation.[30] The results also support the findings of Manski and Pepper (2000), who question the validity of the extremely high returns to schooling obtained in many studies that use instrumental variables techniques.[31]

The effect of a GED on wages ranges from .01% to 5.5% across the five occupations, and the effect is statistically significant in every occupation except craftsmen. Note that the monetary return to earning a GED is less than the monetary return to completing one year of high school in every occupation. In contrast, Cameron and Heckman (1993) find that the GED does not have a positive effect on wages using a regression which assumes that earning a GED is exogenous. At the other extreme, 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 $10-19\%$.

The estimates of the firm and occupation specific human capital parameters are presented in the bottom half of Table B1. These parameters measure the change in log wages accruing to workers as their firm specific capital increases. For example, moving to the second firm specific human capital level increases a professional's wage by approximately 12%, and moving to the third level results in an additional increase of 5.9%.[32] The relationship between firm specific capital and wages is concave for professionals, sales, and service workers, and convex for craftsmen (level 2: 4.4%, level

[30] The model estimated by Belzil and Hansen (B&H) shares the basic methodology used in this study, as both studies estimate a dynamic programming model of education and earnings, but there are many differences between the models. A few of the larger differences are: 1)B&H focus on education so they do not model occupational choices, 2) school interuption is exogenous in B&H, while it is endogenous in the present model, 3) B&H abstract away from firm and occupation specific capital and job matching since their focus is on education, 4) B&H use a more flexible spline function specification of the returns to education.

[31] Manski and Pepper (2000) use a monotone IV assumption to determine that the upper bound on the increase in log-wages from completing four years of college is .397. In this paper, the estimated returns to completing college range from .388 for professionals and managers to .124 for operatives and laborers.

[32] The first human capital level is set equal to zero since it is not separately identifed from the constant in the wage equation.

3: 11.2%). The importance of firm specific capital varies across occupations, with operatives and laborers realizing the lowest wage increases with firm tenure (9.5% at level 3), and service workers realizing the largest gains (24.8% at level 3). Across all occupations the probability of firm specific skill increase is essentially equal to one, so wages increase quickly with firm tenure for two years before levelling out.[33]

The importance of occupation specific capital varies widely across occupations. Both operatives and laborers and sales and clerical workers realize essentially no gain from occupation specific capital, and service workers realize a relatively modest gain of 3.9% when their occupation specific skills reach the highest level. In contrast, professional and managerial workers realize a wage gain of 17% at the third level occupation specific capital, while craftsmen experience a wage gain of 13% at the third level. The relationship between wages and occupation specific capital is convex for professionals and managers, since moving to the second occupation specific capital level increases wages by only 2%, while moving to the third level increases wages by an additional 15%. In contrast, craftsmen realize a large wage gain of 9.6% when moving to the second occupation specific capital level, but moving to the next level increases wages by only an additional 3.6%. In addition, the probability of occupation specific skill increase is substantially lower for craftsmen compared to professionals (.46 vs. .77).

One important consideration is the extent to which the discrete levels of firm and occupation specific human capital are able to capture the patterns in wage growth found in the NLSY. Most of the skill increase probabilities are very close to one, with the exception of the rates of increase for professional and craftsmen occupation specific human capital. When skill increase probabilities are close to one, wages will increase early in jobs but the highest level of human capital will be reached quickly. The concern is that the discrete levels approach will understate on-the-job wage

[33] Rapid wage growth with firm tenure early in jobs that subsides at higher levels of tenure has been found in several studies. For example, Altonji and Shakotko (1987) find that the first year of tenure increases wages by 11%. Dustman and Meghir (2005) report returns to firm tenure for unskilled German workers of 4% per year during the first 5 years of tenure, but the returns are zero for higher levels of tenure.

growth. Appendix C compares OLS estimates of a quadratic specification of a simple wage equation to one that uses three discrete levels of human capital. This Appendix shows that the fit of the quadratic and discrete levels wage equations are virtually identical, with $R^2(quadratic) = .3063$ and $R^2(levels) = .3007$. A more detailed discussion is presented in Appendix C, but it appears that modelling human capital using a discrete number of human capital levels performs extremely well relative to the commonly estimated quadratic functional form, and does not lead to a serious underestimate of the importance of firm and occupation specific human capital.

The estimates of the standard deviations of the random wage shock (σ_e) and pecuniary job match value (σ_ψ) show that both job matching and random wage shocks play an important role in determining wages, and suggest that mobility between firms provides the opportunity for substantial wage increases. Table 6 quantifies the monetary gains to job search (moving to a higher ψ_{ij}) relative to the gains from firm and occupation specific human capital accumulation. The first row of Table 6 shows the percent increase in wages in each occupation accruing to a worker who reaches the highest levels of both firm and occupation specific human capital, while the bottom row depicts the wage gains from moving to higher percentiles of the job match distribution. The potential wage increase from the combination of firm and occupation specific capital varies widely across occupations, ranging from a low of 10% for operatives and laborers to a high of 42% for professionals and managers. There are also substantial gains to job search: a worker who is able to move from the 25th to 75th percentile of the match value distribution realizes a wage gain from job search of 45% ($\exp(.186 - [-.186]) - 1 = .45$), while a worker moving from the 5th to 95th percentile experiences a wage increase of 147%. These results indicate that both human capital accumulation and job search play important roles in determining wage growth over the career, but the relative importance of each effect varies by occupation. The primary source of wage growth for operatives and laborers and sales and clerical workers is finding a good firm match, while in the other occupations the wage gains from human capital accumulation are quite large relative to the

Table 6
Combined Returns to Firm & Occupation-Specific Capital vs. Gains from Job Search

	Professional & Managers	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*
Potential cumulative wage increase from firm & occupation-specific capital	42%	27%	10%	13%	29%
Potential wage gains from job search					
25th percentile match to 75th	45%				
5th percentile match to 95th	147%				

Notes: Gains to firm and occupation-specific human capital are computed using the human capital level parameter estimates (potential wage increase = exp(firm HC level 3+ occ. HC level 3)-1).
Gains to job search are based on the percentiles of the pecuniary job match value (ψ) distribution.

Table 7
Simulated Choice Frequencies by Endowment Type

	Type 1		*Type 2*		*Type 3*		*Type 4*	
Choice percentages at age 21								
Attending school	4.48%		72.99%		7.47%		15.84%	
Unemployed	13.22%		10.02%		13.07%		72.98%	
Professional & managerial	5.54%		11.09%		8.76%		.93%	
Craftsmen	24.31%		2.86%		18.25%		2.48%	
Operatives & laborers	29.85%		5.55%		29.17%		3.11%	
Sales & Clerical	12.58%		18.25%		13.07%		4.04%	
Service	12.15%		11.99%		15.23%		2.17%	
Choice percentages at age 27								
Attending school	.64%		14.49%		.43%		.93%	
Unemployed	2.56%		2.68%		3.02%		64.91%	
Professional & managerial	28.78%		56.17%		30.75%		5.90%	
Craftsmen	36.25%		1.07%		28.59%		4.97%	
Operatives & laborers	20.04%		3.22%		27.01%		8.07%	
Sales & clerical	7.25%		27.37%		4.60%		9.32%	
Service	5.12%		8.05%		6.03%		6.21%	
Value functions & wages at age 27	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**
Value function of optimal choice at age 27	41.86	7.82	67.89	8.91	43.01	8.12	13.01	5.38
Wage at age 27	9.92	.42	9.89	.39	9.42	.40	9.44	.42

Notes: Based on a simulation of 2,000 people. Average simulated wages are conditional on employment.
Value functions represent the discounted expected value of lifetime utility.

potential gains from job search.

5.3 Career Choices & Heterogeneity in Skills and Preferences

Table B2 which is located in Appendix B presents the estimates of the log-wage equation intercepts (μ's) and non-pecuniary utility flow intercepts (ϕ's) for each of the four types of people in the model, along with the estimated proportion of each type in the population. These parameters are all statistically different from zero at the 5% level.[34] 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 wage intercepts show that there is substantial variation in market ability both across and within types. Type 1's have the highest ability in each occupation, while type 2's have the second highest ability in all occupations except service. Differences in the log wage intercepts correspond approximately to percentage changes in wages, so a person's endowment type greatly influences their expected earnings in each occupation. For example, holding the effects of all state variables constant, a type 1 person's expected wage in the sales and clerical occupation is approximately 37% higher than a type 2's expected wage, 44% higher than a type 3's expected wage, and 56% higher than a type 4's expected wage. The across type standard deviations of the μ's show that professional and managerial ability varies the most across people, while the service occupation has the least dispersion in ability.

The non-pecuniary intercepts (ϕ's) 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 non-wage employment intercepts are negative across all occupations and types, which indicates that people experience disutility from employment relative to leisure. The non-wage employment intercepts vary widely across occupations, which indicates that there is

[34] It is common for t-statistics in nonlinear structural models to be very large. See, for example Rust (1987), Berkovec and Stern (1991), or Keane and Wolpin (1997).

substantial heterogeneity in preferences for employment in different occupations across people.

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 6.074 log yearly wage units for type 1's, to a high of 16.778 for type 2's. The disaggregation of ability into market skills (μ's) and school ability or preference (ϕ^s) shows that the two dimensions of ability are far from perfectly positively correlated. Type 1's have the highest market ability in each occupation but the lowest schooling ability.

Table 7 quantifies the impact of heterogeneity in skills and preferences on career outcomes by summarizing career choices for each endowment type based on simulated data generated from the structural model. At age 21 there are already substantial differences in career outcomes across types. Almost 73% of the highest schooling ability people, type 2's, are attending school at age 21. In contrast, the majority of type 1 and 2's have finished attending school and are working in blue collar occupations as craftsmen or operatives and laborers. Type 4's, who experience the highest disutility from working and also have the lowest endowment of market ability have a 73% unemployment rate at age 21. At age 27 types have specialized in different types of employment as a result of variation in skills and preferences. Type 2's are essentially white collar workers, since 56% are employed as professionals and managers, and 27% are employed as sales and clerical workers. Simulated choices at age 27 for types 1 and 3 are quite similar: 29% vs. 31% professional, 36% vs. 29% craftsmen, and 20% vs. 27% laborers.

It is clear that occupational and educational choices are strongly impacted by heterogeneity in skills and preferences, but it is not obvious how this heterogeneity affects key career outcomes such as wages, and, more importantly, total utility. The final section of Table 7 addresses these questions by showing the mean simulated value functions along with mean accepted wages for each type at age 27. Differences in the simulated value functions across types show how the discounted expected value of lifetime utility is impacted by heterogeneity in ability and preferences. The discounted expected value of lifetime utility at age 27 for a type 2 worker is approximately 1.5 times higher than a type 1 or type 3 worker. The discounted expected value of lifetime utility at age 27 for a type 2 worker is 5 times higher than a type 4 worker. Type 4 workers on average spend a large portion of their careers unemployed due to both low market skills and high employment disutility. Differences in skill endowments and preferences have a substantial impact on welfare at age 27. Interestingly, type 3's are better off on average than type 1's (43.01 vs. 41.86), even though type 1's have higher market ability in each occupation and receive a lower disutility from working in the three occupations most often chosen by type 1's and type 3's at age 27.

It is somewhat surprising that type 2's have the highest average discounted expected utility flow at age 27 even though their expected wages based on market ability are over 30% lower than a type 1 person. The simulated data shows that the large gap in market ability between type 1's and type 2's does not translate into inequality wages at age 27, since average log wages for these types are virtually identical at 9.92 for type 1's compared to 9.89 for type 2's. Gaps in market ability (μ's) do not necessarily translate into wage inequality because pursuit of comparative advantage drives the observed wage gap below the gap in ability. Most importantly, type 2's have a large advantage in schooling ability, and as a result they complete more years of schooling than type 1's, as shown in the large gap in school attendance at age 21 between these two types. Most type 2's work as professionals and managers where completing college raises wages by 39% over what a high school educated worker would earn. The combined effect of differences in schooling ability, differences in employment preferences, and pursuit of comparative advantage result in a wage gap well below the market ability gap between type 1's and 2's.

The variation in discounted expected lifetime utility across types suggests that skill and preference heterogeneity is an important determinant of welfare inequality. A regression of the discounted expected value of lifetime utility on type dummy variables explains 64% of the variation in lifetime utility across people, so heterogeneity in skills and preferences is a key determinant of welfare. One implication of this result is that job search models that do not incorporate occupations are missing a key determinant of welfare. The remaining 36% of variation in utility is caused by random shocks to wages and non-pecuniary utility flows, the arrival of job matches, and randomness in human capital improvement. To put this result in context, Keane and Wolpin (1997) find that heterogeneity in schooling ability and market ability explains 90% of the variation in lifetime utility. The addition of job search, firm specific capital, and random shocks to non-pecuniary utility to an occupational choice model reduces the importance of permanent heterogeneity in determining welfare, but its impact is still substantial.[35]

Another 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 in the model. The results of this decomposition are presented in Table 8. The top half of Table 8 shows the percentage of the unexplained variation in wages in each occupation 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 8 shows that 31% of the unexplained variation in wages in the professional and managerial occupation is due to variation in skill endowments (μ^q) across people, and 31% 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 31% to 40%. These results indicate that job matching is a significant determinant of wages in all five occupations. In contrast, the importance of occupation

[35] In addition to the previously stated differences between Keane and Wolpin (1997) and the present model, other key differences that may impact the importance of permanent heterogeneity are the level of aggregation of civilian occupations (five compared to two in K+W), the exclusion of military employment from the present model, and the inclusion of heterogeneity in employment preferences along with heterogeneity in ability in the present model.

Table 8
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	31%	31%	38%
Craftsmen	15%	38%	47%
Operatives & Laborers	15%	38%	47%
Sales & Clerical	19%	36%	45%
Service	10%	40%	50%
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	48%	0%	52%
Craftsmen	28%	0%	72%
Operatives & Laborers	21%	0%	79%
Sales & Clerical	23%	0%	77%
Service	26%	0%	74%
School	53%	0%	47%

Table 9: The Impact of Human Capital, Job Matching, and Occupational Matching on Welfare and Wages

Counterfactuals	**Total Log-Wages**		**Total Utility**	
	Total	% Change from baseline	Total	% Change from baseline
Baseline (estimated model)	192,599	0%	143,545	0%
1) Eliminate effect of firm and occupation specific capital on wages	187,452	-2.7%	139,871	-2.5%
2) Eliminate effect of education on wages	184,625	-4.1%	132,380	-7.8%
3) Workers randomly assigned to firms, never allowed to switch firms	155,956	-19%	96,862	-33%
4) Workers randomly assigned to occupations, never allowed to switch occupations	135,437	-30%	122,526	-15%

Notes: Computed using samples of 2,000 simulated people. Total wages and utility are the sums of accepted wages and realized one period utility flows over people and years. See Section 6.1 of the text for a description of the restrictions imposed under each counterfactual.

specific skills varies more widely across occupations. The fraction of unexplained variation in wages attributed to skill heterogeneity ranges from a high of 31% for professionals and managers to a low of 10% for service workers.

The bottom half of Table 8 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. Variation in preferences across people is most important in determining utility for professional and managerial workers, and is least important for operatives and laborers. The firm specific non-pecuniary match value (ξ) accounts for approximately zero percent of the variation in non-pecuniary utility flows across all five occupations. This result indicates that the permanent firm specific non-pecuniary match value is not an important determinant of employment utility. However, it does not imply that non-wage considerations are of minor importance when workers are searching for jobs since the random non-pecuniary utility shock (ε) is an important determinant of utility across all five occupations. The extremely small estimate of the standard deviation of the non-pecuniary match value (σ_ξ) only implies that randomness in non-wage utility is not correlated over time within a firm after accounting for correlation at the level of occupations, which is captured by ϕ. There are several possible interpretations of this result. It is possible that non-pecuniary job characteristics vary substantially from one year to the next for a given firm. Alternatively, even if non-pecuniary conditions are constant over time within a firm, a worker may evaluate these working conditions and fringe benefits differently from one year to the next.

5.4 Non-pecuniary Utility Flows

The coefficients of the deterministic portions of the non-pecuniary utility flow equations are presented in Tables B3 and B4. The coefficients are interpreted as changes in utility flows relative to the base choice of unemployment. For example, each year of high school completed increases the one-period utility flow from attending school by .569 relative to the value of being unemployed. The estimates of the switching costs show that workers incur a mobility cost of approximately 2.66 when switching firms or re-entering employment from unemployment. Incurring the moving cost has the same effect on utility as a 93% decrease in wages, so these switching costs are quite large.[36]

The parameter estimates in Table B4 show that the effects of observable variables on employment non-pecuniary utility vary widely across occupations. For example, each additional year of education increases employment utility by .773 for professional and managerial workers, but each year of education decreases employment utility by .649 for craftsmen. Employment utility increases sharply in each occupation as workers accumulate both firm and occupation specific capital. One interpretation of these effects is that acquiring greater skills makes a job easier, which reduces the disutility of working. The positive effect of firm tenure on non-pecuniary utility may also arise from fringe benefits that increase with firm tenure.

6 Counterfactual Experiments

One of the major advantages of structural estimation relative to reduced form approaches is that structural models can be used to conduct counterfactual experiments to examine the impact of changes in the economic environment on behavior and welfare. This section uses the structural model to conduct counterfactual simulations that quantify the effects of changes in the economic environment on lifetime earnings and utility. The first set of counterfactual experiments examines the contributions of human capital, job matching, and occupational matching to the welfare and wages of workers. The second counterfactual quantifies the pecuniary and non-pecuniary costs of

[36] See Berkovec and Stern (1991) or Lee and Wolpin (2005) for examples of other dynamic structural models with large estimated switching costs.

job displacement.

6.1 The Social Value of Human Capital, Job Matching, and Occupational Matching

The first row of Table 9 shows the total log-wages earned and utility realized by workers in 2,000 simulated careers generated from the structural model. This baseline simulation is based on the model as specified in Section 3 along with the simulated maximum likelihood parameter estimates. Comparing the baseline simulation to simulations that implement counterfactual changes in the model provides information about the effects of human capital, job search, and occupational matching on total earnings (log-wages) and welfare (total utility). The first counterfactual examines the impact of firm and occupation specific human capital on wages and utility by eliminating the wage effects of these types of human capital, calculating the value functions under this restriction, and then using the new value functions to simulate career choices. The effects of firm and occupation specific human capital on wages are eliminated by setting the following wage equation parameters equal to zero: $\beta_j^q = 0$, $q = 1, ..., 5$, $j = 8, ..., 13$. Eliminating the effects of firm and occupation specific capital on wages decreases total earnings by 2.7%, while the total utility realized by workers in the simulated economy decreases by 2.5%. In other words, the total pecuniary value of firm and occupation specific capital is 2.7% of earnings, while the social value is 2.5% of total utility. The counterfactuals measure the net effect of each change, which includes many offsetting behavioral effects. For example, one effect of eliminating the returns to firm and occupation specific capital is to decrease wages because this change eliminates on the job wage growth. This effect is offset to some degree by the fact that eliminating on the job wage growth reduces moving costs in the form of human capital that is lost when workers switch firms or occupations.

The second counterfactual quantifies the impact of education on wages by showing how wages and total welfare would change if the pecuniary returns to education were eliminated. This restriction is imposed by setting the effects of high school and college education on wages equal to zero across all occupations ($\beta_3^q = 0$, $\beta_4^q = 0$, $q = 1, ..., 5$). The results of this counterfactual, shown in Table 9, reveal that the combined pecuniary value of high school and college education is 4.1% of total earnings, while the total social value is 7.8% of total utility. This counterfactual simulation captures the net effect of eliminating the returns to education, where the wage losses from the reduction in human capital are offset to some extent because a decrease in the payoff to attending school increases the number of years worked by the average person in the simulated sample. The social value of education is larger than the pecuniary value because when the pecuniary return to education is eliminated people choose to accumulate less schooling, which decreases non-pecuniary utility because there is a consumption value to attending school and because education increases the employment non-pecuniary utility flow in many occupations.

The third and fourth counterfactuals shown in Table 9 examine the pecuniary and social gains to matching between workers and firms and workers and occupations. The benefits to society resulting from job search are quantified in the third counterfactual, where workers are randomly matched to firms and not allowed to switch firms during their career. In this world, the gains to job search are eliminated because workers are unable to search for jobs across firms. However, workers are free to self select into their optimal occupation. This counterfactual shows that eliminating job search reduces total earnings by 19%. The social value of job search is even larger than the monetary gains: eliminating job search decreases total welfare by 33%. Note that the value of job search to society dwarfs the social value of human capital. The combined total social value of education and firm and occupation specific capital is approximately one-third as large as the value of job search (10% of total utility vs. 33%).

The social value of workers self selecting into occupations (and switching occupations) is captured in the fourth counterfactual, where each worker is randomly matched to an occupation for his entire career. This counterfactual eliminates occupational mobility as well as self selection in occupational choices based on abilities and preferences, but workers are free to move between firms over the course of their career. Randomly assigning workers to occupations reduces total earnings by 30%, so there are substantial monetary gains to society from allowing workers to match themselves to occupations based on their skills and preferences. The total welfare gain to society from matching between workers and occupations is equal to 15% of total utility.

The counterfactual experiments presented in this section highlight the social gains arising from the mobility of workers across firms and occupations as they make optimal career decisions. Although the parameters indicate that there are substantial pecuniary returns to occupation and firm specific human capital, the counterfactual simulations show that job search and self selection into occupations are far more important determinants of wages and total welfare. The large gains arising from mobility between firms and occupations suggest that it is crucial to incorporate both job search and occupational choices when studying labor market dynamics since they are both key determinants of total earnings and overall welfare.

6.2 The Welfare Impact of Job Displacement

The structural parameter estimates highlight the importance of human capital, gains from job search, and non-pecuniary utility in determining career outcomes and welfare. Given the importance of these effects, job displacement may result in large pecuniary and non-pecuniary costs by destroying both human capital and productive job matches. The structural model separately identifies the effects of wages and non-pecuniary utility in determining total utility, so the model is well suited to quantify both the wage losses caused by displacement as well as the overall impact on welfare. Previous research on the cost of displacement has focused on the monetary cost of job loss, which ignores the potentially large role of non-pecuniary utility in determining welfare.

Table 10 shows the impact of job displacement at age 25 on wages and discounted expected utility flows for workers in each occupation. The table shows that a job displacement at age 25 decreases the discounted expected value of lifetime utility by $17\% - 27\%$ depending on a worker's

Table 10: The Impact of Job Displacement at Age 25 on Wages and Utility

Occupation at age 24		Age 25			Age 26			Age 30		
	Averages	Baseline simulation	Displacement simulation	Change	Baseline	Displacement simulation	Change	Baseline	Displacement simulation	Change
Professional & Managers	Wage	9.80	0	-9.8	9.87	9.73	-.14	9.93	9.92	-.01
	Value function	66.04	53.39	-19%	64.34	54.18	-16%	52.29	48.99	-6.3%
Craftsmen	Wage	9.68	0	-9.68	9.72	9.58	-.14	9.70	9.72	.02
	Value function	51.17	37.48	-27%	47.26	38.18	-19%	27.88	26.56	-4.7%
Operatives & laborers	Wage	9.47	0	-9.47	9.49	9.41	-.08	9.62	9.60	-.02
	Value function	49.85	37.45	-25%	44.15	37.36	-15%	27.65	27.11	-1.9%
Sales & clerical	Wage	9.75	0	-9.75	9.77	9.62	-.15	9.85	9.87	.02
	Value function	58.80	45.52	-23%	53.90	46.18	-14%	43.78	43.34	-1%
Service	Wage	9.47	0	-9.47	9.56	9.43	-.13	9.75	9.75	0
	Value function	52.32	43.27	-17%	48.88	42.33	-13%	34.94	35.57	-1.7%

Notes: Baseline and displaced simulations are computed using samples of 2,000 simulated people. In the displacement simulation workers become unemployed at age 25. Average wages are conditional on employment, and the wage is the log of the yearly wage. Average value functions represent the discounted expected value of lifetime utility.

occupation at the time of displacement, with craftsmen experiencing the largest losses, and service workers experiencing the smallest losses. In the year following a displacement average log wages for workers who have found a new job are between .08 and .15 lower than their expected wages in a world where the job loss had not occurred. The negative impact of displacement on wages dissipates over time as workers accumulate human capital and find new job matches. Five years after displacement (age 30), the wages of displaced workers in each occupation are approximately equal to the wages of non-displaced workers.[37] Although the wage effect has subsided after five years, the effect on total utility has not, since the average discounted expected value of lifetime utility for a displaced worker at age 30 is still $1\% - 6.3\%$ lower than a non-displaced worker.

7 Conclusion

This paper formulates and structurally estimates a dynamic model of educational attainment, occupational choices, and job search that incorporates self-selection in occupational and educational choices, endogenous accumulation of firm and occupation specific human capital, and job search based on firm level wage and non-pecuniary matching. The model integrates the dynamic human capital occupational choice framework developed by Keane and Wolpin (1997) with the job search approach to labor market dynamics. The benefit of estimating a model that nests both of these approaches to analyzing career choices is that the parameter estimates provide direct evidence about the relative importance of features of human capital models relative to features of job search models in explaining the determination of wages and total utility over the career.

The parameter estimates show that wages increase with both firm and occupation specific capital, and that the human capital wage function varies widely across occupations. The potential total wage gains from firm and occupation specific capital range from a low of 10% for operatives and laborers to a high of 42% for professionals and managers. While the wage gains from human

[37] Existing evidence on the long term impact of job displacement is mixed. Jacoson, LaLonde, and Sullivan (1993) find that in their non-mass layoff sample wages recover 5 years after displacement, but Ruhm (1991) finds evidence of persistent earnings reductions.

capital are substantial, the estimates of the job search portion of the model indicate that mobility to better job matches is also a key source of wage growth. In addition, heterogeneity in occupation specific ability, school ability, and preferences for employment in different occupations is a powerful determinant of career choices and overall welfare. This heterogeneity accounts for approximately 64% of the variation in discounted expected lifetime utility across people.

The structural model is used to conduct counterfactual simulations that quantify the contributions of human capital accumulation, job search, and occupational matching to total income and overall welfare. These simulations reveal that eliminating the pecuniary returns to firm and occupation specific human capital would reduce total welfare by 2.5%, eliminating occupational matching would reduce welfare by 15%, and eliminating the gains to firm matching would reduce welfare by 33%. These results indicate that the importance of labor mobility in determining overall welfare far exceeds the importance of human capital. Workers realize large gains as they make optimal occupational choices and inter-firm mobility decisions, which implies that policies that promote worker mobility by lowering mobility costs or search frictions have the potential to increase wages and welfare by promoting the efficient assignment of workers to firms and occupations.

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 (2006). 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.

Appendix B: Structural Parameter Estimates

Table B1
Structural Model Estimates

Variable	Professional & managers	Craftsmen	Operatives & laborers	Sales & clerical	Service
	Occupations				
Log Wage Equation:					
Age (β_1)	-.018	.096	.003	.036	-.011
	(.0007)	(.0092)	(.0008)	(.0018)	(.0012)
Age^2/100 (β_2)	.089	-.408	.036	-.036	.205
	(.0031)	(.0153)	(.0032)	(.0044)	(.0058)
Years of high school (β_3)	.048	.013	.059	.029	.020
	(.0061)	(.0007)	(.0063)	(.0008)	(.0008)
Years of college (β_4)	.097	.046	.031	.073	.097
	(.0032)	(.0044)	(.0041)	(.0019)	(.0024)
Age ≤ 17 (β_5)	-.272	-.069	-.201	-.180	-.032
	(.0033)	(.0039)	(.0130)	(.0133)	(.0047)
18 ≤ Age ≤ 21 (β_6)	-.272	-.036	-.165	-.193	-.042
	(.0014)	(.0020)	(.0203)	(.0023)	(.0026)
GED (β_7)	.020	.001	.055	.021	.011
	(.0015)	(.0019)	(.0033)	(.0026)	(.0020)
Firm-specific HC: level 1 (β_8)	0.00&	0.00&	0.00&	0.00&	0.00&
Firm-specific HC: level 2 (β_9)	.120	.044	.042	.082	.168
	(.0048)	(.0031)	(.0021)	(.0017)	(.0019)
Firm-specific HC: level 3 (β_{10})	.179	.112	.095	.125	.248
	(.0202)	(.0190)	(.0267)	(.0292)	(.0552)
Occupation-specific HC: level 1 (β_{11})	0.00&	0.00&	0.00&	0.00&	0.00&
Occupation-specific HC: level 2 (β_{12})	.020	.096	.0006	.00001	.033
	(.0013)	(.0033)	(.0003)	(.00002)	(.0018)
Occupation-specific HC: level 3 (β_{13})	.173	.132	.0006	.00001	.039
	(.0348)	(.0320)	(.0004)	(.00003)	(.0371)
Probability that firm-specific human capital increases (λ_f)	.998	.999	.999	.999	.999
	(.0193)	(.0388)	(.0194)	(.056)	(.0468)
Probability that occupation-specific human capital increases (λ_f)	.777	.463	.999	.190	.999
	(.0179)	(.0238)	(.0487)	(.0361)	(.0511)
Error Standard Deviations	Estimate	Stan. Error			
True randomness in wages (σ_e)	.306	.0080			
Non-Pecuniary firm match value (σ_ξ)	.00046	.0025			
Pecuniary firm match value (σ_ψ)	.275	.0131			
Extreme value parameter (τ)	3.234	.2107			

Notes: & indicates the parameter is fixed at the stated value, not estimated. Standard errors in parentheses.
Age is measured as true age minus 15.

Table B2
Structural Model Estimates

Variable	*Type 1*	*Type 2*	*Type 3*	*Type 4*	*Across types*	
Log-wage Intercepts					Mean	Std dev
Professional & managerial (μ^1)	9.677 (.047)	9.244 (.046)	8.974 (.0552)	9.014 (.0639)	9.215	.276
Craftsmen (μ^2)	9.108 (.048)	8.878 (.0596)	8.660 (.0477)	8.735 (.0566)	8.834	.173
Operatives & laborers (μ^3)	9.346 (.047)	9.011 (.0393)	8.990 (.0576)	8.827 (.0398)	9.050	.175
Sales & clerical (μ^4)	9.319 (.052)	8.947 (.0574)	8.873 (.0595)	8.762 (.0562)	8.976	.198
Service (μ^5)	9.162 (.062)	8.841 (.0658)	8.836 (.0615)	8.854 (.0769)	8.916	.135
Within-type mean	9.322	8.984	8.866	8.838		
Within-type std dev	.222	.159	.132	.109		
Non-pecuniary Intercepts						
Professional & managerial (ϕ^1)	-28.351 (.1386)	-25.340 (.1251)	-27.563 (.1446)	-37.208 (.1731)	-28.856	4.004
Craftsmen (ϕ^2)	-21.335 (.1120)	-23.825 (.1309)	-21.005 (.1123)	-28.065 (.1113)	-23.051	2.575
Operatives & laborers (ϕ^3)	-16.20 (.0865)	-14.541 (.0828)	-15.531 (.0672)	-20.905 (.0937)	-16.370	2.170
Sales & clerical (ϕ^4)	-22.862 (.2076)	-19.902 (.0938)	-23.003 (.1524)	-26.975 (.1362)	-22.859	2.290
Service (ϕ^5)	-19.345 (.0880)	-16.427 (.0834)	-18.804 (.0913)	-23.955 (.1098)	-19.214	2.452
School (ϕ^s)	6.074 (.0451)	16.778 (.1720)	6.848 (.0489)	7.440 (.0592)	9.353	4.423
Within-type mean of (ϕ^1-ϕ^5)	-21.618	-20.007	-21.181	-27.422		
Within-type std dev (ϕ^1-ϕ^5)	4.513	4.629	4.518	6.141		
Type Probabilities	.233 (.053)	.260 (.038)	.332 (.054)	.175 (.033)		

Note: Standard errors are in parentheses.

Table B3
Structural Model Estimates

Variable	*Estimate*	*Standard Error*
Discount factor (δ)	.95[&]	
School Utility Flow		
Age (γ_{s1})	-3.666	.0757
Age^2/100 (γ_{s2})	9.591	.1712
Attending college (γ_{s3})	.671	.0179
Attending graduate school (γ_{s4})	-2.264	.0459
Years of high school (γ_{s5})	.569	.0168
Years of college (γ_{s6})	.488	.0611
School While Employed Utility Flow		
Age (γ_{sw1})	-5.271	.0894
Age^2/100 (γ_{sw2})	24.74	.9561
Years of high school (γ_{sw3})	4.138	.0583
Years of college (γ_{sw4})	1.054	.0098
GED Utility Flow		
Constant (γ_{g1})	-.950	.0068
Age (γ_{g2})	-10.409	.8946
Switching Costs		
Cost of moving to a new firm (firm to firm transitions) (α_{10})	2.661	.0377
School re-entry cost (γ_{s7})	-2.376	.0721
Cost of moving to a new job from non-employment (α_9)	2.658	.0407
Costs of Working while Attending School		
Work in high school (γ_{sw6})	6.497	.2385
Work in college (γ_{sw7})	11.548	.3794
Work in graduate school (γ_{sw8})	12.093	.4292
Within-firm Job Offer Probabilities		
Offer from professional & managerial (π_1)	.249	.0144
Offer from craftsmen (π_2)	.215	.0160
Offer from operatives & laborers (π_3)	.226	.0155
Offer from sales & clerical (π_4)	.224	.0198
Offer from service (π_5)	.085	.0075

Notes: & indicates the parameter is fixed at the stated value, not estimated.

Table B4
Structural Model Estimates

Variable	*Occupations* Professional & Managers	Craftsmen	Operatives & Laborers	Sales & Clerical	Service
Employment Non-Pecuniary Utility Flows:					
Age (α_1)	1.927 (.126)	2.035 (.1373)	.860 (.1036)	1.761 (.1518)	.850 (.1720)
$Age^2/100$ (α_2)	-7.995 (.5662)	-10.098 (.6794)	-4.105 (.4842)	-10.689 (.8518)	-4.028 (.5637)
Education (α_3)	.773 (.066)	-.649 (.0499)	-.620 (.0458)	.248 (.0403)	.024 (.0541)
Occupation-Specific human capital (α_4)	5.532 (.256)	3.657 (.2348)	2.524 (.3172)	2.217 (.4416)	1.984 (.3138)
Firm-Specific human capital (α_5)	2.028 (.0162)	2.508 (.0701)	2.072 (.0615)	2.564 (.0786)	2.405 (.0627)
High school diploma (α_6)	.639 (.1135)	2.222 (.1894)	1.749 (.1877)	1.862 (.1763)	.756 (.1592)
College diploma (α_7)	2.492 (.2103)	4.803 (.3765)	4.319 (.3743)	5.127 (.3290)	3.527 (.2891)
GED (α_8)	1.422 (.2490)	1.718 (.2189)	2.335 (.2234)	1.711 (.1933)	2.982 (.2072)
Log-likelihood	-15,422				

Notes: Standard errors in parentheses.

Appendix C: The Wage Equation – Discrete Levels vs. Quadratic Specifications

Table C1: Quadratic vs. Discrete Level Log-Wage Equations

Variable	Quadratic	Discrete Levels
Firm tenure	.0498 (.0076)	---
Firm tenure^2	-.0021 (.0009)	---
Occupation experience	.0202 (.0071)	---
Occupation experience^2	-.0003 (.0006)	---
Firm tenure level 2	---	.0910 (.0141)
Firm tenure level 3	---	.0872 (.0201)
Occupation experience level 2	---	.0129 (.0116)
Occupation experience level 3	---	.0643 (.0212)
Observations	8,297	8,297
R^2	.3063	.3007

Notes: Both regression equations are estimated using OLS. Standard errors are adjusted for clustering. Tenure and experience are measured in years. The other explanatory variables included are identical to those in the structural log wage equation. The level dummy variables are defined as: Firm tenure level 1 = 1 if firm tenure>0, and Firm tenure level 2 = 1 if firm tenure>1, Occupation experience level 1 = 1 if occupation experience>0, and Occupation experience level 2 = 1 if occupation experience>1.

Both regressions are estimated using OLS. Firm tenure and occupation experience are endogenous, so these relationships should not be interpreted as causal. The intent of this analysis is only to provide a simple descriptive analysis of the observed patterns in wage growth. The results show that the R^2's of the quadratic and discrete level specifications are extremely close, so both approaches provide approximately the same fit to the wage data observed in the NLSY. The two specifications perform equally well because in the NLSY data, most wage growth occurs early in jobs. It is important to remember that this analysis considers young men at the start of their career (ages 16-30), so average firm tenure and occupation experience are only 2.2 years and 2.4 years, respectively. Given this feature of the data, it is perhaps not surprising that the discrete levels approach performs so well.

Another closely related point is that the discrete levels approach used in the OLS regressions is much more restrictive that the approach used in the structural model. In the structural model the rate of skill increase is not restricted during estimation, so that conditional on working for one year, it is possible for skills to rarely increase, or always increase. However, in the OLS equation the rate of increase is restricted to equal one, which means that workers always move to the next level after each year of work. The fact that many of the increase parameters are close to one in the estimated structural model does not signal an identification problem, it simply means that given three discrete levels, the best way to fit the wage data is by having skills increase quickly. Given the strong performance of the OLS discrete levels specification, it is not surprising that many of the estimated structural skill increase probabilities are close to one.

References

[1] Altonji, Joseph and Robert Shakotko (1987). “Do Wages Rise with Job Seniority?" *The Review of Economic Studies*, v. 54, no. 3.

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

[3] Belzil, Christian and Jorgen Hansen (2002). “Unobserved Ability and the Return to Schooling." *Econometrica*, v. 70, no. 5.

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

[5] 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.

[6] Blau, David (1991). “Search for Nonwage Job Characteristics: A Test of the Reservation Wage Hypothesis.” *Journal of Labor Economics,* v. 9, no. 2: 186-205.

[7] Brien, Michael, Lee Lillard, and Steven Stern (2006). “Cohabitation, Marriage and Divorce in a Model of Match Quality.” *International Economic Review,* v. 47, no. 2.

[8] Brown, Meta and Christopher Flinn (2004) “Investment in Child Quality Over Marital States.” Working Paper.

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

[10] Cameron, Stephen, and Christopher Taber (2003). “Estimation of Educational Borrowing Constraints using Returns to Schooling." Working Paper.

[11] Canals-Cerda, Jose, and Steven Stern (2002). “Empirical Search Models.” in *Search Theory and Unemployment*, eds. Steven Woodbury and Carl Davidson, Kluwer Academic Publications.

[12] Card, David. (1999) “The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, v. 3, eds. Orley Ashenfelter and David Card. Elsevier, Amsterdam.

[13] Dey, Matthew and Christopher Flinn (2005). “An Equilibrium Model of Health Insurance Provision and Wage Determination." *Econometrica*, v. 73, no. 2.

[14] Dustmann, Christian and Costas Meghir (2005). “Wages, Experience and Seniority." *Review of Economic Studies*, v. 72, no. 1.

[15] Eckstein, Zvi and Gerard van den Berg (2006). “Empirical Labor Search: A Survey." *Journal of Econometrics*, forthcoming.

[16] 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.

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

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

[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] Hwang, Hae-shin, Dale Mortensen, and Robert Reed (1998). “Hedonic Wages and Labor Market Search." *Journal of Labor Economics*, v. 16, no. 4.‘

[21] Jacobson, Louis, Robert LaLonde and Daniel Sullivan (1993). “Earnings Losses of Displaced Workers.” *American Economic Review*, v. 83 no. 4: 685-709.

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

[23] Kambourov, Gueorgui and Iorii Manovskii (2006). “Occupational Specificity of Human Capital.” Working Paper.

[24] 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.

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

[26] 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.

[27] Lee, Donghoon (2005). “An Estimable Dynamic General Equilibrium Model of School, Work, and Occupational Choice." *International Economic Review*, v. 46: 1-34.

[28] Lee, Donghoon and Kenneth Wolpin (2006). “Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica*, v. 74, no. 1.

[29] Light, Audrey (2001) “In-School Work Experience and the Returns to Schooling.” *Journal of Labor Economics*, v. 19, no. 1: pp. 65-93.

[30] Manski, Charles and John Pepper (2000). “Monotone Instrumental Variables : With an Application to the Returns to Schooling." *Econometrica*, v. 68, no. 4: 997-1010.

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

[32] McFadden, Daniel (1981). “Econometric Models of Probabilistic Choice." In C. Manski and D. McFadden (eds.) *Structural Analysis of Discrete Data with Econometric Applications*, pp. 198-272. MIT Press, Cambridge.

[33] Mellow, Wesley and Hal Sider (1983). “Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics*, v. 1, no.4.

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

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

[36] Pavan, Ronni (2006). "Career Choice and Wage Growth." Working Paper.

[37] Ruhm, Christopher (1991). "Are Workers Permanently Scarred by Job Displacements?" *American Economic Review,* v. 81: 319-324.

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

[39] Rust, John (1997). "Using Randomization to Break the Curse of Dimensionality." *Econometrica*, v. 65, no. 3.

[40] 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.

[41] Stern, Steven (1997). "Simulation Based Estimation." *Journal of Economic Literature*, v. 35, no. 4: 2006-2039.

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

[43] Topel, Robert (1991). "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, v. 99: 145-176.

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

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

[46] 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.