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

Sullivan, Paul

Bureau of Labor Statistics

June 2008

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

Paul Sullivan*
Bureau of Labor Statistics

June 2008

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. Wage growth occurs in the model because workers move between firms and occupations as they search for suitable job matches and because workers endogenously accumulate firm and occupation specific human capital. Simulations performed using the estimated model 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 31%, 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 2.8%.

*I would like to thank John Pepper, William Johnson, and especially Steven Stern for their invaluable assistance during this project. I thank three anonymous referees and the editor, Petra Todd, for helpful comments and suggestions. This research has benefited from discussions with Loren Smith, John Bound, Timothy Erickson, Randal Verbrugge, and Daniel Mackay, and from the comments of numerous seminar participants. I gratefully acknowledge financial support for this research provided by the Bankard Fund for Political Economy at the University of Virginia, the University of Virginia College of Arts and Sciences, and a National Institute on Aging Post Doctoral Fellowship at the University of Michigan (grant number AG00221-14). 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. Email: Sullivan.Paul.Joseph@bls.gov. Phone: (202) 691-6593.

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. These decisions are inherently interrelated, yet existing research has generally examined educational attainment, occupational choices, and on-the-job human capital accumulation separately from decisions about job search.[1] As a result of this separation in the literature, there is currently no way to assess the importance of interactions between these decisions, or to determine the importance of human capital relative to the importance of mobility between firms and occupations in determining wage growth over the career.

The goal of this paper is to address this gap in the literature 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. The model allows workers to accumulate firm and occupation specific human capital as they move between firms and occupations over their careers. Estimating the model provides evidence about the relative importance of human capital, job search, and matching between workers and occupations in determining wages and total utility. The parameter estimates reveal that each aspect of the model is quantitatively important and necessary to understand the evolution of wages over the career. However, the main empirical conclusion that emerges from this analysis is that self selection in occupational choices and mobility between firms are much more important determinants of total earnings and utility than the combined effects of firm and occupation specific human capital.

In the career choice model developed in this paper, 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 impose constraints on the mobility of workers between 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

[1] See Keane and Wolpin (1997), Lee (2005), and Lee and Wolpin (2006) for examples of dynamic human capital models that focus on occupational choices and human capital accumulation. A recent survey of the extensive job search literature is provided by Eckstein and van den Berg (2006).

specific human capital.[2] The value of employment varies over the five occupations in the 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.[3]

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 estimated structural wage equation reveals that expected wage offers tend to increase as workers accumulate firm and occupation specific human capital. Wage offers are also impacted by the quality of the match between a worker and his employer, and by the quality of the match between a worker and his occupation. The estimated structural model is used to perform counterfactual simulations which reveal that eliminating the gains from matching between workers and occupations would reduce total wages by 31%, eliminating the gains from job search would reduce wages by 19%, and eliminating the combined effects of firm and occupation specific human capital on wages would reduce earnings by only 2.8%. Existing research has been unable to determine the importance of each of these effects because the typical approaches to studying wage growth over the career examine the contributions of either human capital accumulation or mobility in isolation, but do not attempt to jointly estimate the importance of each facet of the career decision problem.

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. 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,

[2] Throughout this paper the term human capital is used to refer to wage growth that occurs with tenure in both firms and occupations, since actual human capital is of course unobserved. This paper does not attempt to address the difficult issue of separately identifying human capital effects from other sources of within-job wage growth, such as promotions that may be unrelated to productivity growth, or wage growth due to contracts designed to provide incentives to workers.

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

Lee 2005, Lee and Wolpin 2006).[4] 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.[5]

The model developed in this paper expands on the occupational choice model of Keane and Wolpin (1997) by incorporating 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 aspect of the career decision problem in explaining career choices, wages, and total utility.

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 14-22 years old when first interviewed in 1979. 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.

The NLSY consists of a nationally representative core sample, a military sample, and a supplemen-

[4] See Eckstein and Wolpin (1999) and Belzil and Hansen (2002) for examples of papers that estimate dynamic structural models that abstract away from occupational choices and focus on the endogenous accumulation of education.

[5] Berkovec and Stern (1991) and Wolpin (1992) develop search models that include firm specific capital but these models do not incorporate occupational choices. McCall (1990) and Neal (1999) develop search models that incorporate occupations, but these models do not include human capital accumulation.

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.

tal sample that over-samples blacks, Hispanics, and economically disadvantaged whites. This analysis uses only white men from the nationally representative core sample, and these individuals are followed from age 16 until age 30. 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.[6] The data are aggregated using an approach similar to that of Keane and Wolpin (1997) to assign yearly employment status and school attendance. See Appendix A for a detailed description of the procedures used to aggregate the data.

The NLSY data provides information on occupational codes at the three digit level. However, the cost of estimating the model increases substantially as the number of occupations increases, so using extremely detailed occupational classifications is not feasible. 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 comparable existing research.[7]

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. 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.

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

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

[7] Keane and Wolpin (1997) and Lee (2005) aggregate the data into only two occupations (blue and white collar). Lee and Wolpin (2006) allow workers to choose between blue, white, and pink collar employment in both the service and goods sectors.

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% 54.40%	29.94% 27.38%
22-25	50.09% 47.14%	26.85% 23.39%
26-30	40.76% 37.86%	17.61% 14.83%
All Ages	49.78% 46.56%	24.69% 21.75%

Note: Probabilities are computed using all consecutive years of employment observed in the data for each age group. The top entry of 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*
Professional & Managers	83.28 86.10	4.22 2.84	3.00 2.48	7.35 6.61	2.15 1.97
Craftsmen	7.25 5.40	75.59 77.54	13.05 12.15	2.55 4.36	1.57 .55
Operatives & Laborers	4.74 4.73	14.90 13.53	68.98 71.24	7.66 7.52	3.71 2.98
Sales & Clerical	20.45 17.31	4.60 6.01	10.76 8.87	61.94 65.36	2.25 2.45
Service	10.53 8.82	7.22 7.01	9.32 8.05	4.51 6.23	68.42 69.89

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 of each cell is computed using the NLSY data, and the bottom entry is computed using simulated data generated using the estimated structural model.

as attending school. 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. Keane and Wolpin (1997) report a similarly high rate of unemployment using slightly different definitions of employment and school attendance.

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.[8] 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%.

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. 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.

3 Economic Model of Career Choices

Each individual's career is modeled as a finite horizon, discrete time dynamic programming problem. 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. Each period, an

[8] 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.

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.

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 $s = \textit{school}$, $g = \textit{GED}$ and $u = \textit{unemployed}$.[9] 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,

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

[9] 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 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. The permanent worker-firm productivity match is represented by ψ_{ij}. 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.

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

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

The first term in brackets represents the influence of the state vector on non-pecuniary utility flows and is discussed in more detail in the following paragraph. The second term in brackets captures the effect of person specific heterogeneity in preferences for attending school (ϕ_i^s), being unemployed (ϕ_i^u), and being employed in occupation q (ϕ_i^q). The non-pecuniary occupation match value, ϕ_i^q, represents the random component of person i's preference for working in occupation q. 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 final term, ε_{ikt}, is a shock to the non-pecuniary utility that person i receives from choice combination k at time t.

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

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

The term in brackets contains the occupation and firm specific non-pecuniary utility flows. The occupation specific portion of this flow, $\theta_q(S_{it})$, is a function of the state vector that is allowed to vary over occupations. 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. The second line of equation 4 contains the schooling cost function for attending school while not employed ($C^s(S_{it})$) and employed ($C^{sw}(S_{it})$). 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.[10] Any dual

[10] 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}). This framework is adopted in many papers such as

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

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

Next, consider the feasible choices for a person employed in occupation q. At the start of period t the individual receives one new job offer from a firm in each of the five occupations and has the option to attend school, earn a GED, or become unemployed. In addition, an employed individual always has the option of remaining at his current firm and staying in his current occupation (q). Job offers from new occupations at the current firm are received randomly, where workers receive either zero or one such offer per year. Let π_j denote the probability that a worker receives an offer to work in occupation j at his current firm, where $j \neq q$. Let π_{nq} be the probability that a worker employed in occupation q does not receive an offer to switch occupations within his current firm.

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

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

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. Let a_t represent an individual's age. Educational

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. Given the available data these two models are observationally equivalent, see Eckstein and van den Berg (2006) for a detailed discussion.

attainment is summarized by the number of years of high school and college completed, h_t and c_t, and a dummy variable indicating whether or not a GED has been earned, g_t. Work experience is captured by the amount of firm specific human capital (f_t) and occupation specific human capital (o_t) in the occupation that the person worked in most recently. Let $O_t \in [1, 2, ..., 5]$ indicate the occupation in which a person was most recently employed. Let L_t be a variable that indicates a person's previous choice, where $L_t = \{1, ..., 5\}$ refers to working in occupations one through five, $L_t = 6$ indicates attending school full time, and $L_t = 7$ indicates unemployment.

Given this notation, the state vector is $S_t = \{a_t, h_t, c_t, g_t, f_t, o_t, O_t, L_t, \xi_t, \psi_t\}$. 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. 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

$$\begin{aligned} f_t &\in FC = \{f(1), ..., f(P)\} \\ o_t &\in OC = \{o(1), ..., o(P)\}. \end{aligned}$$

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.[11] There are separate skill increase probabilities for firm and occupation specific capital, and the rates of skill increase are also allowed to vary across occupations. The skill increase parameters are $\{\lambda_f^k, \lambda_o^k, \quad k = 1, ..., 5\}$, where the subscripts

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

f and o refer to firm and occupation specific capital, and k indexes occupations. The human capital transition probabilities are known by agents in the model. 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.

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. Viewing increases in human capital as a stochastic event is consistent with the idea that years of work experience only serve as a proxy for an individuals unobservable true 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 deterministic components of the utility function and his endowment of market skills and occupation specific non-pecuniary match values. The maximization problem can be represented in terms of alternative specific value functions.[12] The value function for an individual with discount factor δ employed in occupation q is

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

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, since human capital evolves stochastically. Consider the first summation in equation 7. Each term in

[12] The value function is a function of S_t, but this argument is suppressed for brevity of notation.

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.

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

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

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

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

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

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

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.[13]

[13] Pavan (2006) estimates a model of career choices that focuses on modeling mobility at the very disaggregated level of three-digit occupations and industries. In contrast, the model developed in this paper uses much more highly aggregated occupation groups, but allows for endogenous educational attainment and heterogeneous human capital effects across occupations.

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 using standard backwards recursion techniques. 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 so a discrete approximation to these distributions is used when solving the optimization problem. Assume that the random choice-specific utility shocks are distributed extreme value, with distribution function $F(\varepsilon) = \exp\{-\exp(-\frac{\varepsilon}{\tau})\}$, and with variance $\pi^2\tau^2/6$. The assumption that ε is 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 and preferences. Using an approach similar to Heckman and Singer (1984) and Keane and Wolpin (1997), 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 $\chi_m(h_{i1})$ as the proportion of the mth type in the population, where the argument h_{i1} indicates that the type probabilities are conditioned on the number of years of high school that an individual has completed as of age 16 (individuals reach age 16 in time period $t = 1$ by definition).[14] 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

[14] Following Keane and Wolpin (1997) type probabilities are allowed to vary between individuals who have not completed the 10th grade by age 16 and those who have complete at least the 10th grade by age 16.

of skills and preferences allows for a wide range of patterns of comparative advantages in skills and heterogeneity in preferences.

3.3.2 Calculating the Value Functions

The major complication arises from the fact that as the model is specified the Emax integrals do not have closed form solutions. This paper uses an interpolation algorithm that builds on the one developed by Keane and Wolpin (1994) to decrease the amount of time needed to solve the optimization problem. The details of the simulation and interpolation solution method are presented in Appendix B.

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.

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

$$\begin{aligned} w_q(S_{it}) \quad = \quad & \beta_1^q a_{it} + \beta_2^q a_{it}^2/100 + \beta_3^q h_{it} + \beta_4^q c_{it} + \beta_5^q 1[a_{it} \leq 17] + \\ & \beta_6^q 1[a_{it} \geq 18 \cap a_{it} \leq 21] + \beta_7^q g_{it} \\ & + \beta_8^q 1[f_{it} = f(1)] + \beta_9^q 1[f_{it} = f(2)] + \beta_{10}^q 1[f_{it} = f(3)] \\ & + \beta_{11}^q 1[o_{it} = o(1)] + \beta_{12}^q 1[o_{it} = o(2)] + \beta_{13}^q 1[o_{it} = o(3)]. \end{aligned} \tag{12}$$

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.

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

$$\begin{aligned}\theta_q(S_{it}) \quad &= \quad \alpha_1^q a_{it} + \alpha_2^q a_{it}^2/100 + \alpha_3^q(h_{it} + c_{it}) + \alpha_4^q o_{it} + \alpha_5^q f_{it} + \alpha_6^q hd_{it} \\ &\quad + \alpha_7^q cd_{it} + \alpha_8^q g_{it} + \alpha_9^q 1[L_{it} > 5] + \alpha_{10}^q NF_{it} \qquad q = 1, ..., 5.\end{aligned} \tag{13}$$

The inclusion of explanatory variables in the employment non-pecuniary utility flow equations allows observable variables to have a direct impact on employment utility in addition to any 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. The cost function for attending school is

$$\begin{aligned}c^S(S_{it}) \quad &= \quad \gamma_{s1}a_{it} + \gamma_{s2}a_{it}^2/100 + \gamma_{s3}hd_{it} + \gamma_{s4}cd_{it} + \gamma_{s5}h_{it} + \gamma_{s6}c_{it} + \gamma_{s7}1[L_{it} \neq 6] \\ c^{SW}(S_{it}) \quad &= \quad \gamma_{sw1}a_{it} + \gamma_{sw2}a_{it}^2/100 + \gamma_{sw3}h_{it} + \gamma_{sw4}c_{it} + \gamma_{s7}1[L_{it} \neq 6] \\ &\quad + \gamma_{sw6}(h_{it} \leq 4) + \gamma_{sw7}(h_{it} = 4 \cap c_{it} \leq 4) + \gamma_{sw8}(c_{it} \geq 4).\end{aligned} \tag{14}$$

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.[15]

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

$$c^g(S_{it}) = \gamma_{g1} + \gamma_{g2}a_{it}. \tag{15}$$

Within-firm job offer probabilities are specified as multinomial logit, so the probability of receiving a

[15] 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.

job offer from occupation j at the current firm is

$$\pi_j = \frac{\exp(\rho_j)}{\sum_{k=1}^{5} \exp(\rho_k)}. \tag{16}$$

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.[16]

4.2 The Likelihood Function

The solution to the dynamic programming problem provides the choice specific value functions which are used in the construction of the likelihood function. Let Θ represent the parameters of the structural model. 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.

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

$$\begin{aligned} L_i(\Theta \mid \Phi_i = \Phi_k) \quad = \quad & \int \cdots \int [\int \int \left(\prod_{t=1}^{\widetilde{T}_i} \Pr[O_{it} \mid \Omega, S_{it}, o_{it}, f_{it}, \Phi_i = \Phi_k) \right) \\ & dF(o_i) dF(f_i)] dF(\Omega). \end{aligned} \tag{17}$$

Note that the path probability for each person is integrated over the distributions of occupation and firm specific human capital (o_i and f_i) because these variables are unobserved. The likelihood contribution is also integrated over the joint distribution of $\Omega = \{\psi, \xi, e\}$, because these match values and choice specific utility shocks are not observed.

The high dimensional integrals in the likelihood function are approximated using simulation methods. The details of the simulation algorithm along with a derivation of the outcome probabilities are provided in Appendix C. Let $L_i^S(\Theta \mid \Phi_i = \Phi_m)$ represent the simulated type-specific likelihood contribution for person i. The simulated likelihood function for the sample is the product over the N people

[16] 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.

in the sample of a weighted average of the type-specific simulated likelihoods, where the weights are the type probabilities ($\chi_m(h_{i1})$),

$$L^S(\Theta) = \prod_{i=1}^{N} \sum_{m=1}^{M} \chi_m(h_{i1}) L_i^S(\Theta \mid \Phi_i = \Phi_m). \tag{18}$$

The vector of parameters $\widehat{\Theta}$ that maximizes equation number 18 is the simulated maximum likelihood estimate of Θ.

Standard errors are computed using a parametric bootstrap estimator of the covariance matrix of $\widehat{\Theta}$. This approach to estimating standard errors has been successfully applied in complex structural models such as the one estimated by Engers and Stern (2002). The bootstrapped standard errors are computed by using the parameter estimates $\widehat{\Theta}$ to simulate R samples of data, and then re-estimating $\widehat{\Theta}$ using each simulated sample. The parameter estimates from the R simulated samples are used to construct an estimate of the variance of the parameter vector. This procedure is extremely computer intensive because the model has nearly 200 parameters that must be re-estimated for each simulated sample. Also, recall that each likelihood evaluation is quite expensive because it involves solving the dynamic programming problem. Given these considerations, the standard errors are estimated using $R = 35$ simulated datasets.[17]

4.3 Identification

Although the career choice model is fairly complex and contains a large number of parameters it is still fairly straightforward to provide some intuition for how the parameters of the model are identified. The goal is to estimate the parameters of occupation specific wage offer equations along with parameters of non-pecuniary utility flow equations. In many respects this situation is analogous to simple linear estimation because the data contain information about the correlations between observable variables and discrete choices and wages. Similarly, the parameters of the stochastic processes that allow for

[17] The computational burden of the parametric bootstrap may be lessened by taking k steps of a derivative based optimization routine when estimating Θ for each simulated sample instead of allowing the optimization routine to continue until convergence in each sample. In this work, experimentation showed that $k = 4$ provides a very close approximation to the value of $\widehat{\Theta}$ that would be obtained if the number of optimization steps was not restricted. See Davidson and MacKinnon (1999) for a detailed discussion of this $k-$step parametric bootstrap.

serially correlated unobservables within firms and occupations are identified by the panel nature of the data.

Static sample selection models are also informative about identification of the structural model. The data contain information about an individual's wages and occupational choices, but of course wages are only observed for an individual's chosen occupation. The solution to the agent's optimization problem provides the sample selection rules that are used to estimate a selection corrected wage equation for each occupation. The obvious analogy is to static selection models that are estimated by maximum likelihood or by two-step procedures, and the major difference between the static and dynamic models is that the selection rules in the dynamic model are provided by the numerical solution of the agent's optimization problem.

Identification of the parameters of the non-pecuniary utility flow equations follows from the fact that the data contain information about discrete career choices along with wages. It is possible to estimate the effects of observable variables on non-pecuniary utility because the data provide information about the extent to which individuals' career choices are not completely explained by variation in wages across occupations. To the extent that observed variables are correlated with observed choices after conditioning on wages, this provides information about the impact of the observed variables on non-pecuniary utility. For example, if college educated workers work as professionals more often then one would expect solely based on occupational wage differentials, this suggests that the professional non-pecuniary utility flow is increasing in years of completed education.

5 Structural Parameter Estimates

Table 6, Panels A-D 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 estimated structural model is used to simulate a sample of $4,000$ 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 NLSY and simulated data. The discrepancies between simulated and actual mean log wages range from zero to .08 across occupations, and the model also matches the standard deviations of wages in each occupation 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. 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 choice proportions disaggregated by age for both the NLSY data and simulated data. The model qualitatively 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.

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	9.78	9.59	9.38	9.54	9.33
Wage std dev: NLSY data	.54	.45	.45	.51	.47
Wage std dev: simulated data	.51	.48	.45	.50	.47

Note: Simulated wages computed from a sample of 4,000 people. Yearly wages are in logs.

Figure 1: Actual & Simulated Mean Log-wages

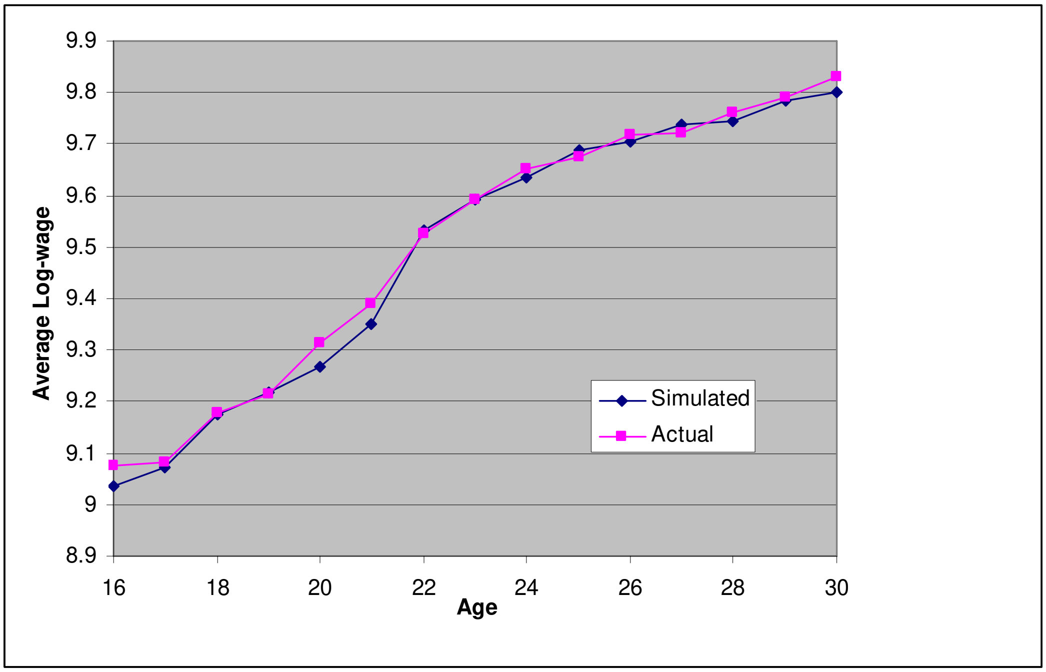


Table 7
Combined Returns to Firm & Occupation-Specific Capital vs. Gains from Job Search

	Professional & Managers	*Craftsmen*	*Operatives & Laborers*	*Sales & Clerical*	*Service*
Potential cumulative wage increase from firm & occupation-specific capital	42%	27%	10%	13%	34%
Potential wage gains from job search					
25^{th} percentile match to 75^{th}	45%				
5th percentile match to 95^{th}	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.

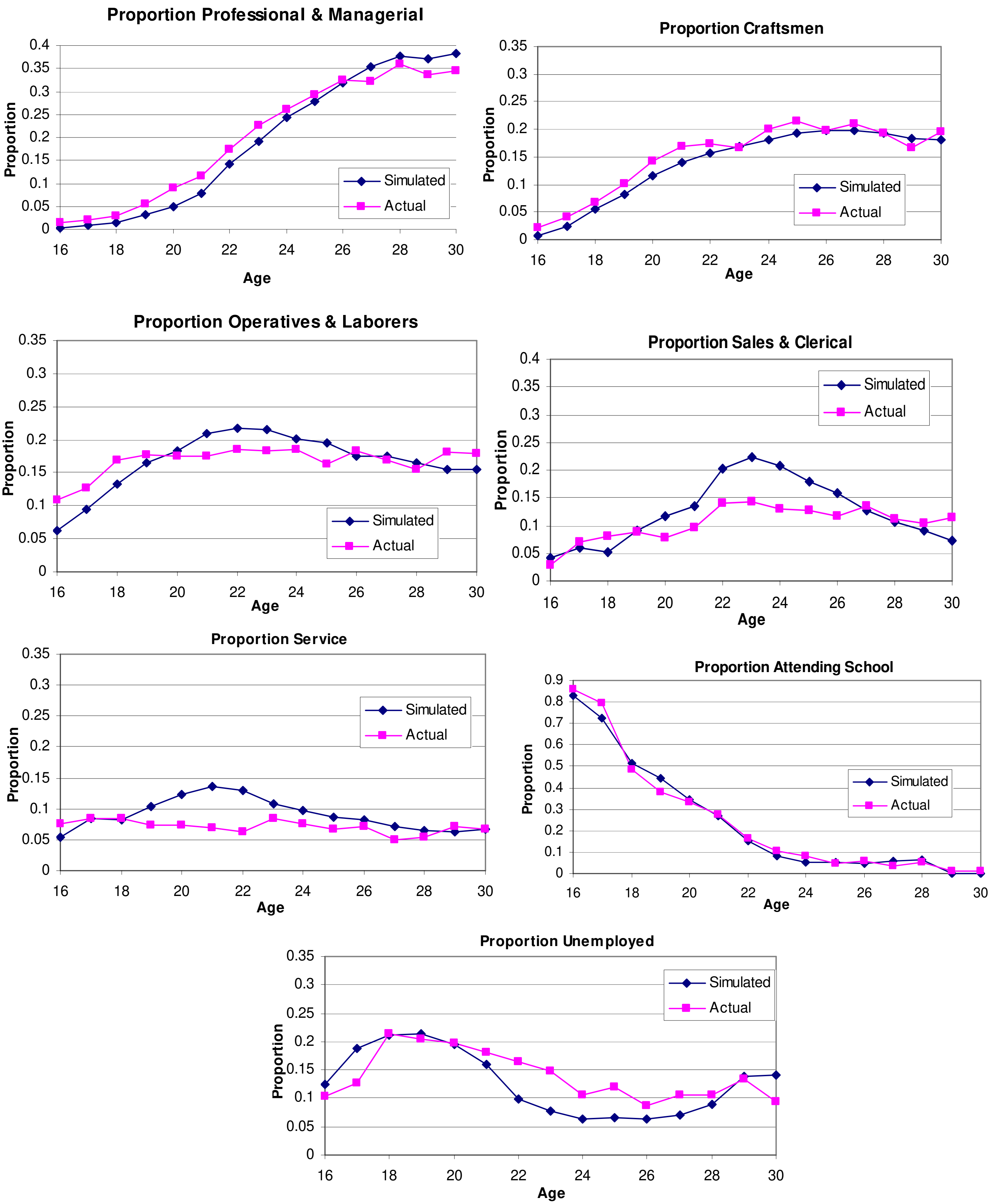
The model tracks the downward age trend in school attendance extremely closely. The simulated data reproduces the general qualitative age pattern in unemployment found in the NLSY data, although the model under predicts the unemployment rate in the early to mid twenties. The model also overstates employment in the sales and clerical occupation during the mid twenties.

5.2 The Log Wage Equation: Human Capital & Job Search

The estimates of the log wage equation parameters found in Table 6, Panel A 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 1.4% for craftsmen to a high of 5.6% for operatives and laborers. Interestingly, the null hypothesis that the effect of completing a year of high school has no effect on wages is not rejected at the 5% level for the craftsmen, sales and clerical, and service occupations. The monetary return to attending college also varies widely across occupations, and is statistically different from zero in all occupations. Completing a year of college increases wages by approximately 9.2% for professional and managerial workers, while a year of college increases wages by only 3.2% for an operative or laborer. 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 7.3% for professional and managerial workers, 3.2% for craftsmen, 4.2% for operatives and laborers, 5.4% for sales and clerical workers, and 6.7% for service workers. These results are consistent with the relatively low average

Figure 2
Choice Proportions by Age: Actual and Simulated Data



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.[18] 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.[19]

The point estimate of the effect of a GED on wages ranges from .10% to 5.6% across the five occupations, although the effect is not statistically different from zero in any occupation. These results are consistent with those of Cameron and Heckman (1993), who 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 6, Panel A. 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%. The relationship between firm specific capital and wages is concave for professionals, sales, and service workers, and convex for craftsmen (level 2: 4.1%, level 3: 10.9%). The importance of firm specific capital varies widely across occupations, with operatives and laborers realizing the lowest wage increases with firm tenure (9.7% at level 3), and service workers realizing the largest gains (25.4% 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.[20] During estimation these probabilities converged to numbers that were essentially equal to one,

[18] The model estimated by Belzil and Hansen (2002) (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.

[19] 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 .39. In this paper, the estimated returns to completing college range from .37 for professionals and managers to .13 for operatives and laborers.

[20] Rapid wage growth with firm tenure early in jobs that subsides at higher levels of tenure has been found in several

Table 6: Panel A
Structural Model Estimates

Variable	*Professional & managers*	*Craftsmen*	*Operatives & laborers*	*Sales & clerical*	*Service*
	Occupations				
Log Wage Equation:					
Age (β_1)	-.019	.098	.003	.036	-.010
	(.011)	(.011)	(.008)	(.015)	(.010)
$Age^2/100$ (β_2)	.085	-.406	.036	-.037	.206
	(.046)	(.050)	(.037)	(.071)	(.051)
Years of high school (β_3)	.048	.014	.056	.029	.021
	(.016)	(.011)	(.009)	(.020)	(.012)
Years of college (β_4)	.092	.047	.032	.072	.103
	(.007)	(.008)	(.008)	(.005)	(.009)
Age $\leq$ 17 (β_5)	-.272	-.069	-.196	-.180	-.032
	(.066)	(.058)	(.033)	(.055)	(.036)
18 $\leq$ Age $\leq$ 21 (β_6)	-.270	-.036	-.162	-.194	-.042
	(.022)	(.019)	(.015)	(.021)	(.018)
GED (β_7)	.021	.001	.056	.021	.011
	(.037)	(.047)	(.042)	(.043)	(.036)
Firm-specific HC: level 1 (β_8)	.000[&]	.000[&]	.000[&]	.000[&]	.000[&]
Firm-specific HC: level 2 (β_9)	.119	.041	.044	.081	.157
	(.012)	(.015)	(.012)	(.014)	(.023)
Firm-specific HC: level 3 (β_{10})	.179	.109	.097	.124	.254
	(.011)	(.015)	(.015)	(.020)	(.021)
Occupation-specific HC: level 1 (β_{11})	.000[&]	.000[&]	.000[&]	.000[&]	.000[&]
Occupation-specific HC: level 2 (β_{12})	.024	.092	.000	.000	.046
	(.020)	(.016)	(---)	(---)	(.015)
Occupation-specific HC: level 3 (β_{13})	.172	.130	.000	.000	.046
	(.018)	(.026)	(---)	(---)	(.015)
Probability that firm-specific human capital increases (λ_f)	.999 (---)	.999 (---)	.999 (---)	.999 (---)	.999 (---)
Probability that occupation-specific human capital increases (λ_f)	.777 (.061)	.463 (.018)	.999 (---)	.189 (.040)	.999 (---)
Error Standard Deviations	Estimate	Stan. Error			
True randomness in wages (σ_e)	.309	0.001			
Non-Pecuniary firm match value (σ_ξ)	.000	---			
Pecuniary firm match value (σ_ψ)	.276	0.004			
Extreme value parameter (τ)	3.29	0.189			
Log-likelihood	-15,252				

Notes: & indicates the parameter is fixed at the stated value and not estimated because it is not identified. Standard errors are in parentheses. (---) denotes parameters which were fixed during estimation at the stated value, so standard errors are not reported. Age is measured as true age minus 15.

so these parameters were fixed at the stated value during estimation. For this reason, bootstrapped standard errors are not reported for these parameters.

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 4.6% 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 2.4%, while moving to the third level increases wages by nearly an additional 15%. In contrast, craftsmen realize a large wage gain of 9.2% when moving to the second occupation specific capital level, but moving to the next level increases wages by only an additional 3.8%. In addition, the probability of occupation specific skill increase is substantially lower for craftsmen compared to professionals (.46 vs. .77).[21]

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, so 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 growth. Unfortunately, keeping track of years of human capital is not feasible given that the state space of the model is already very large. One way of addressing this concern is by comparing OLS estimates of a quadratic specification of a simple wage equation to one that uses three discrete levels to capture the effects of firm and occupation specific human capital. These specifications of the wage equation provide virtually the same fit to the data, with $R^2(quadratic) = .3063$ and $R^2(levels) = .3007$, and both specifications contain the same number of

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.

[21] See Kambourov and Manovskii (2007) for estimates of the return to occupation tenure under the assumption that the returns to human capital are identical across all occupations.

parameters. 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.[22]

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 7 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 7 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% ($\exp(.186 - [-.186]) - 1 = .45$). 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 potential gains from job search.

5.3 Career Choices & Heterogeneity in Skills and Preferences

Table 6, Panel B presents the estimates of the log-wage equation intercepts and non-pecuniary utility flow intercepts for each of the four types of people in the model, along with the estimated proportion of each type in the population.

[22] It is important to remember that this analysis focuses on young men at the start of their career (ages 16-30), so average firm tenure and occupation tenure are only 2.2 and 2.4 years. Given this feature of the data, it is perhaps not surprising that the discrete levels approach performs so well.

The log wage intercepts represent skill endowments in each of the five occupations. 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 approximately 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 55% higher than a type 4's expected wage. Across occupations, professional and managerial ability varies the most in the population (standard deviation=.27), while the service occupation has the least dispersion in ability (standard deviation=.13).

The non-pecuniary intercepts reflect a person's preferences for working in each occupation and attending school. These parameters are measured in log yearly wage units relative to the base choice of unemployment. The 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 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 6.07 log yearly wage units for type 1's, to a high of 16.77 for type 2's. The disaggregation of ability into market skills and school ability or preference 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 8 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.

Table 6: Panel B
Structural Model Estimates

Variable	***Type 1***	***Type 2***	***Type 3***	***Type 4***
Log-wage Intercepts				
Professional & managerial (μ^1)	9.68 (.057)	9.24 (.053)	8.97 (.055)	9.01 (.061)
Craftsmen (μ^2)	9.11 (.075)	8.88 (.082)	8.66 (.075)	8.73 (.02)
Operatives & laborers (μ^3)	9.35 (.051)	9.00 (.046)	8.99 (.046)	8.82 (.049)
Sales & clerical (μ^4)	9.32 (.110)	8.95 (.110)	8.87 (.112)	8.76 (.113)
Service (μ^5)	9.16 (.063)	8.84 (.068)	8.84 (.058)	8.85 (.068)
Non-pecuniary Intercepts				
Professional & managerial (ϕ^1)	-28.35 (1.60)	-25.34 (1.35)	-27.56 (1.55)	-37.21 (2.06)
Craftsmen (ϕ^2)	-21.33 (1.00)	-23.82 (1.14)	-21.00 (.95)	-28.06 (1.31)
Operatives & laborers (ϕ^3)	-16.20 (.79)	-14.54 (.79)	-15.53 (.83)	-20.90 (1.16)
Sales & clerical (ϕ^4)	-22.86 (1.01)	-19.90 (.97)	-23.00 (1.08)	-26.97 (1.27)
Service (ϕ^5)	-19.34 (.85)	-16.42 (.79)	-18.80 (.81)	-23.95 (1.04)
School (ϕ^S)	6.07 (.63)	16.77 (1.30)	6.85 (.68)	7.44 (.74)
Type Probabilities				
Initial schooling > 9 years	.301 (.025)	.330 (.046)	.331 (.021)	.038 (.002)
Initial schooling ≤ 9 years	.218 (.016)	.144 (.049)	.474 (.029)	.163 (.010)

Note: Standard errors are in parentheses.

Table 6: Panel C
Structural Model Estimates

Variable	Estimate	Variable	Estimate
Discount factor (δ)	.95[&]	***Switching Costs***	
School Utility Flow		Firm to firm transitions (α_{10})	-2.61
			(0.27)
Age (γ_{s1})	-3.68	School re-entry (γ_{s7})	-2.38
	(.33)		(0.30)
Age^2/100 (γ_{s2})	9.59	New job from non-employment (α_9)	-2.66
	(1.41)		(0.27)
Attending college (γ_{s3})	.66	***Costs of Working while Attending School***	
	(.55)		
Attending graduate school (γ_{s4})	-2.26	Work in high school (γ_{sw6})	6.50
	(.69)		(0.62)
Years of high school (γ_{s5})	.56	Work in college (γ_{sw7})	11.55
	(.14)		(0.79)
Years of college (γ_{s6})	.49	Work in graduate school (γ_{sw8})	12.09
	(.12)		(0.94)
School While Employed Utility Flow		***Within-firm Job Offer Probabilities***	
Age (γ_{sw1})	-5.27	Offer from professional & managerial (π_1)	0.25
	(.31)		(0.01)
Age^2/100 (γ_{sw2})	24.75	Offer from craftsmen (π_2)	0.21
	(1.50)		(0.01)
Years of high school (γ_{sw3})	4.15	Offer from operatives & laborers (π_3)	0.23
	(.29)		(0.01)
Years of college (γ_{sw4})	1.07	Offer from sales & clerical (π_4)	0.23
	(.17)		(0.01)
		Offer from service (π_5)	0.09
			(0.01)

Table 6: Panel D
Structural Model Estimates

	Occupations				
Variable	***Professional & Managers***	***Craftsmen***	***Operatives & Laborers***	***Sales & Clerical***	***Service***
Employment Non-Pecuniary Utility Flows:					
Age (α_1)	1.92	2.04	0.86	1.76	0.86
	(.24)	(.19)	(.14)	(.17)	(.14)
Age^2/100 (α_2)	-8.02	-10.10	-4.11	-10.69	-4.03
	(.96)	(1.02)	(.66)	(1.06)	(.72)
Education (α_3)	0.81	-0.65	-0.62	0.26	0.02
	(.13)	(.12)	(.14)	(.16)	(.12)
Occupation-Specific HC (α_4)	5.53	3.66	2.53	2.22	2.08
	(.37)	(.305)	(.21)	(.22)	(.22)
Firm-Specific HC (α_5)	2.03	2.52	2.08	2.56	2.42
	(.18)	(.22)	(.17)	(.17)	(.23)
High school diploma (α_6)	0.62	2.26	1.75	1.86	0.74
	(.43)	(.41)	(.34)	(.42)	(.34)
College diploma (α_7)	2.49	4.82	4.32	5.15	3.54
	(.33)	(.64)	(.46)	(.58)	(.56)
GED (α_8)	1.43	1.72	2.34	1.73	2.99
	(.56)	(.68)	(.45)	(.48)	(.49)

Notes: Standard errors in parentheses.

Approximately 75% 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 77% 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 28% are employed as sales and clerical workers.

The final section of Table 8 shows how heterogeneity impacts wages and utility by showing the mean simulated value functions along with mean accepted wages for each type at age 27. 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, and 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.[23]

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 56% 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 44% 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

[23] Bayes' rule can be used to calculate conditional type probabilities for each person in the sample and see if they are related to family background characteristics. The results of this exercise are available at http://www-personal.umich.edu/~paulsull/supplement.pdf. As in Keane and Wolpin (1997), there is a strong positive correlation between parental income and education levels and the probability that an individual is a white collar, high education type.

Table 8: Simulated Choice Frequencies by Endowment Type

	Type 1		*Type 2*		*Type 3*		*Type 4*	
Choice percentages at age 21								
Attending school	4.48%		75.54%		6.29%		12.91%	
Unemployed	12.85%		8.84%		13.41%		77.51%	
Professional & managerial	6.72%		12.63%		7.66%		0.00%	
Craftsmen	22.29%		2.64%		20.62%		2.39%	
Operatives & laborers	27.95%		7.35%		30.20%		3.35%	
Sales & Clerical	11.79%		19.40%		11.13%		2.39%	
Service	16.86%		11.02%		14.78%		2.87%	
Choice percentages at age 27								
Attending school	.79%		16.15%		1.71%		1.37%	
Unemployed	3.01%		2.00%		3.41%		59.59%	
Professional & managerial	28.48%		56.15%		26.46%		2.74%	
Craftsmen	32.91%		1.23%		27.68%		7.53%	
Operatives & laborers	22.78%		3.08%		28.90%		13.70%	
Sales & clerical	6.49%		28.48%		4.88%		10.27%	
Service	6.33%		8.31%		8.66%		5.48%	
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	43.66	7.94	70.23	8.92	44.88	7.96	14.04	5.49
Wage at age 27	9.95	.42	9.92	.42	9.45	.40	9.42	.47

Notes: Based on a simulation of 4,000 people. Average simulated wages are conditional on employment.

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)	265,321	---	210,623	---
1) Eliminate effect of firm and occupation specific capital on wages	257,860	-2.8%	205,479	-2.4%
2) Eliminate effect of education on wages	255,317	-3.8%	201,413	-4.4%
3) Workers randomly assigned to firms, never allowed to switch firms	213,844	-19%	138,568	-34%
4) Workers randomly assigned to occupations, never allowed to switch occupations	183,030	-31%	176,478	-16%

Notes: Computed using samples of 4,000 simulated people. Total wages and utility are the sums of accepted wages and realized one period utility flows over people and years. See Section 6.2 of the text for a description of the restrictions imposed under each counterfactual.

still substantial.[24]

6 Counterfactual Experiments

The first set of counterfactuals examines the contributions of human capital, job matching, and occupational matching to the wages and total utility of workers.[25]

6.1 A Restricted Model

Before presenting the counterfactuals that quantify the importance of matching between workers and occupations, it is useful to begin by estimating a restricted model that rules out heterogeneity in workers' occupation specific abilities and preferences. This restriction is imposed by estimating the model under the restriction that there is only one type of person, so all workers have identical abilities (μ) and preferences (ϕ). When the null version of the model is estimated, the value of the log-likelihood function is $-19,347$. The likelihood ratio test statistic for the null hypothesis of homogeneity in occupation specific skills and preferences is $8,190$, so the null hypothesis of homogeneity is rejected at any conventional significance level. The large decrease in the log-likelihood function when unobserved heterogeneity is eliminated shows that this feature of the model is necessary to match occupational choices and career outcomes.

6.2 The 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 4,000 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

[24] 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. In addition, it is possible that the use of a continuous match value distribution and discrete unobserved heterogeneity distribution could tend to understate the importance of permanent heterogeneity.

[25] It should be noted that these counterfactual simulations are partial equilibrium in nature. These simulations quantify the effect of changing various parameters on career outcomes, holding all other structural parameters in the model constant. Also, one should keep in mind the fact that these results are based on the NLSY79 cohort of young men used to estimate the model.

provides information about the effects of human capital, job search, and occupational matching on total earnings (log-wages) and the welfare of workers (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 2.8%, while the total utility realized by workers in the simulated economy decreases by 2.4%.

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. This counterfactual produces relatively small changes in the baseline choice distribution. The largest effect is found in the proportion of years spent unemployed, which increases by approximately one percentage point. The diagonal elements in the baseline transition matrix shown in Table 4 all decrease by small amounts ranging from -.10 percentage points for laborers to -1.2 percentage points for service workers. The relatively small changes in the choice distribution and transition matrix show that occupational choices are primarily determined by endowments of skills and preferences, not the heterogeneity across occupations in the effects of firm and occupation specific capital on wages.

The second counterfactual quantifies the impact of education on wages by showing how wages and total utility 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 3.8% of total earnings, while the total social value is 4.4% 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 total 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. Interestingly, eliminating the returns to education has relatively small effects on occupational choices. The simulated occupational choice proportions all change by less than one percentage point under this counterfactual, so the differential returns to education across occupations are not a large determinant of sorting across occupations.

The third and fourth counterfactuals shown in Table 9 examine the pecuniary and total gains to matching between workers and firms and workers and occupations. The benefits to workers 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 total value of job search is even larger than the monetary gains: eliminating job search decreases the total utility of workers by 34%. Note that the value of job search to society dwarfs the social value of human capital. The combined total value of education and firm and occupation specific capital is approximately one-third as large as the value of job search (7% of total utility vs. 34%).

The 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 31%, so there are substantial monetary gains to society from allowing workers to match themselves to occupations based on their skills and preferences. The total utility gain that is attributed to workers making optimal occupational choices and occupational mobility decisions is equal to 16% of total utility.

The counterfactual experiments presented in this section quantify the gains arising from the mobility of workers across firms and occupations as they make optimal career decisions. Although the estimated structural wage equation indicates 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 utility. 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 utility.

6.3 Quantifying the Importance of Comparative Advantage

The previously discussed counterfactuals indicate that there is substantial heterogeneity in abilities and preferences across workers. One way of assessing the importance of comparative advantage effects is to examine how an individual's wages and career outcomes are altered when they are forced to choose a specific occupation other than their optimal occupation. Rather than consider all of the possible combinations of optimal and assigned occupations, this section focuses on two counterfactual scenarios. What would happen to professionals and managers if they were forced to work as operatives and laborers? What would happen to operatives and laborers if they were forced to work as professionals and managers?

When all professional and managerial workers are forced to work as operatives and laborers, the average log wage of these workers decreases by .33 from 9.78 to 9.45. When workers are switched across these occupations in the opposite direction, the average wage for laborers assigned to work as professionals decreases from 9.38 as laborers to 9.21 as professionals.[26] This simulation demonstrates

[26] Notice that professionals have an absolute advantage in both high and low skill employment.

that comparative advantage in occupational choices holds: professionals and laborers both do much better in their chosen occupation than when they are forced to switch occupations.

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 developing a model that nests both of these approaches to analyzing career choices is that the estimated model provides 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 estimated structural wage offer equation shows that wages tend to 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. In addition, there is a considerable amount of heterogeneity in occupation specific ability, school ability, and preferences for employment in different occupations. This heterogeneity accounts for approximately 56% of the variation in discounted expected lifetime utility across people.

Although the estimated wage equation shows that the gains from human capital are substantial, it is difficult to determine the importance of occupation and firm specific human capital relative to the importance of job search and occupational matching by simply examining the parameter estimates. As a result, 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 wages by 2.8%, eliminating occupational matching would reduce wages by 31%, and eliminating the gains to firm matching would reduce wages by 19%. These results indicate that the importance of labor mobility in determining wages 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: Data Aggregation

Yearly school attendance is assigned using detailed information on monthly school attendance and grade completion. The methodology used to assign yearly school attendance consists of several steps. First, the amount of education accumulated by each sample member over the sample period is determined. 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.

Yearly employment status is determined using the weekly labor force record. The yearly employment activity is the activity (a specific employer or unemployment) in which the most weeks were spent during the year. The number of weeks spent unemployed and employed full time at each employer are counted for each decision year. Jobs consisting of less than twenty hours of work per week are counted as time spent unemployed. The work activity in which the most weeks were spent during the school year is coded as the yearly labor force activity. 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. 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.

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. One way of assessing the effects of 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.[27]

Appendix B: Model Solution

B1: Simulating Emax. The Emax integrals do not have closed form solutions, so they are approximated using simulation methods. 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 $\Phi_i = \{\mu_i^1, ..., \mu_i^5, \phi_i^1, ..., \phi_i^5, \phi_i^s, \phi_i^u\}$. Conditional on Ω_t and firm and occupation specific human capital (f_t and o_t), the expected value of the maximum has a closed form solution because of the assumption that ε is distributed extreme value,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \mid \Omega_t, \Phi_i, o_t, f_t\} &= \tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t \mid \Omega_t, \Phi_i, o_t, f_t)}{\tau})]) \\ &= \Psi(d_t \,|\Omega_t, \Phi_i, o_t, f_t) \ , \end{aligned} \tag{19}$$

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, f_t and o_t provides the unconditional expected value of the best choice available next period for each endowment type,

$$E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \,|\, \Phi_i\} = \int \int \left[\int \cdots \int \Psi(d_t \,|\Omega_t, \Phi_i, o_t, f_t) f(\Omega_t) d\Omega_t \right] f(f_t) df_t f(o_t) do_t. \tag{20}$$

[27] 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.

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, $R = 40$.[28] 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 20 over the R draws,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon |\ \Phi_i\} \quad = \quad & \frac{1}{R} \sum_{r=1}^{R} \sum_{h=1}^{P} \Pr[f_t = f_t(h) \,|\, f_{t-1}] \sum_{z=1}^{P} \Pr[o_t = o_t(z) \,|\, o_{t-1}] \times \\ & \Psi(d_t \,|\Omega_t^r, \Phi_i, o_t^z, f_t^h). \end{aligned} \tag{21}$$

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

B2: Interpolation. This paper implements a new interpolating regression function that takes advantage of the assumption that the error term ε is distributed extreme value. 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.[29] If the only source of randomness in the model was the error term ε, then the expected value of the maximum would have the closed form solution shown in equation 19. 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),

[28] 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.

[29] 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.

but it suggests the following functional form for the interpolating regression,

$$\begin{aligned} E \max_{d_t \in D_t} \{\bar{V}(d_t) + \varepsilon \} &= \omega_{0t} + \omega_{1t}\tau(\gamma + \ln[\sum_{d_t \in D_t} \exp(\frac{\bar{V}(d_t)}{\tau})]) \\ &= \omega_{0t} + \omega_{1t}\Psi(d_t) \, . \end{aligned} \tag{22}$$

The parameters ω_{0t} and ω_{1t} are estimated by OLS, and allowed to vary over time. During estimation, the value functions are simulated at approximately 1% of the state space and interpolated at the remaining points. Experimentation shows that the actual and interpolated value functions differ by approximately 1% on average.

Appendix C: Evaluating the Likelihood Function

C1: Simulation of the Likelihood Function. The high dimensional integrals in the likelihood function are simulated using R draws from the joint distribution of Ω and Q draws from the joint distribution of occupation and firm specific human capital. The integral over the joint distribution of human capital is simulated using a modified Geweke, Keane, and Hajivassiliou (GHK) algorithm because the joint distribution of human capital is intractably complex. The type-specific simulated likelihood contribution is

$$L_i^S(\Theta \mid \Phi_i = \Phi_k) = \frac{1}{R}\sum_{r=1}^{R}\frac{1}{Q}\sum_{q=1}^{Q}\prod_{t=1}^{\widetilde{T}_i}\Pr[O_{it}^{rq} \mid \Omega_i^r,\, o_{it}^q, f_{it}^q, \Theta, S_{it}, \Phi_i = \Phi_k). \tag{23}$$

C2: Simulation of the Likelihood Function. With the exception of the integrals 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 f_{it} and o_{it} 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

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

Note that the integral is over the joint distribution of f_i and o_i 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 $o_t^r \,|\, o_{t-1}^r$ and $f_t^r \,|\, f_{t-1}^r$.

2. Compute $\Pr[O_{it} \,|\, o_t^r, f_t^r]$.

3. Compute $\Gamma^r = \Gamma^r * \Pr[O_{it} \,|\, o_t^r, f_t^r]$.

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 f_t^r and o_t^r conditional on the previous draw is very straightforward, while drawing from the complete distribution would be very difficult.

C3: Outcome Probabilities. The most straightforward outcome probability found in the likelihood function is the probability of observing a person attending school or being unemployed. 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}. Conditional on the unobserved state variables except ε, the choice probability is of the multinomial logit form,

$$\Pr(d_{it} = s \,|\, \Omega,\, o_{it}, f_{it}, S_{it}, \Phi_i) = \frac{\exp(V_t(s))}{\sum\limits_{k \in D_t^{ne}} \exp(V_t(k))}. \tag{24}$$

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.

References

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

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

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

[4] 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.

[5] 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.

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

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

[8] 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.

[9] Davidson, Russel and James MacKinnon (1999). “Bootstrap Testing in Nonlinear Models." *International Economic Revie*w, v. 40, no. 2.

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

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

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

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

[14] Engers, Maxim and Steven Stern (2002). “Long Term Care and Family Bargaining." *International Economic Review*, v. 43, no. 1.

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

[16] 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.

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

[18] Kambourov, Gueorgui and Iorii Manovskii (2007). "Occupational Specificity of Human Capital." Forthcoming in the *International Economic Review*.

[19] 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.

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

[21] 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.

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

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

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

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

[26] 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.

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

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

[29] 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.

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

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

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

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

[34] 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.