HUMAN CAPITAL, UNCERTAIN WAGE DISTRIBUTIONS, AND OCCUPATIONAL AND EDUCATIONAL CHOICES

BY PETER F. ORAZEM AND J. PETER Mattila

This study develops a model of occupational and educational choices. Individuals are assumed to choose careers based on expected utility maximization, given uncertainty concerning future earnings. The model directly leads to estimable equations relating career choices to the moments of occupational earnings distributions. We report maximum likelihood estimates using a sample of high school graduates from Maryland school districts, 1951–1969. The eight observed career choices respond as expected to the first and second moments of the earnings distribution and to a measure of school quality. Reasonable estimates of human capital production and supply elasticities are obtained for the career choices.

1. INTRODUCTION

On the surface, the assumption that economic agents select an occupation so as to maximize expected lifetime utility might be considered a truism in that any occupation freely selected by a rational agent must be the agent’s best choice ex ante, even if the decision may prove suboptimal ex post. Yet, despite its apparent acceptance in theoretical work, the assumption of occupational choice on the basis of maximizing expected lifetime utility has not been commonly utilized in conducting empirical studies of occupational choice.

The typical theoretical formulation is that individuals select an occupation \( i \) so as to maximize expected utility, \( E[U(x_i, w_i, Y)] \) across all potential occupations, where \( U \) is an indirect utility function, \( x_i \) is a vector of job characteristics in occupation \( i \), \( w_i \) is lifetime earnings in occupation \( i \), and \( Y \) is income independent of occupational choice.\(^2\) Occupational choices may vary across individuals due to differences in tastes (i.e., differences in attitudes toward job characteristics, income or risk), differences in embodied human capital (i.e. differences in the set of feasible wage offers \( w_i \) caused by different levels of skill), or differences in wealth. However, empirical studies of occupational choice have utilized special cases of the expected utility function that do not fully characterize the theoretical model underlying that choice.

The studies utilizing cross-sectional data typically do not consider uncertainty. Occupational choice is determined by a vector of known personal characteristics (e.g. Polachek 1981; Schmidt and Strauss 1975) or by lifetime earnings which are

---

\(^1\) Helpful comments from the referees are gratefully acknowledged. We thank Lance Kaebelcr for able research assistance. Partial funding for this project came from the University of Wisconsin Institute for Research on Poverty and the U.S. Department of Health and Human Services. The views expressed are solely those of the authors.

\(^2\) See Freeman (1975a) for a review of the theory.
a function solely of known personal characteristics (e.g. Boskin 1974). These studies have the advantage that they typically utilize a logit formulation which allows the agent to choose between any number of potential occupations at once. However, the occupational choice decision is undoubtedly subject to uncertainty regarding future earnings, so these studies fail to capture this important feature of the underlying theory of occupational choice.

Most previous attempts to account for uncertainty in occupational choices have utilized time series data. These models have emphasized the decision to maximize expected earnings rather than expected utility, implicitly imposing risk neutrality on agents making occupational choices. In Richard Freeman’s (1971, 1975a, 1975b) pioneering work, agents make their occupational choice by comparing current wages in an occupation to wages in an alternate occupation. Forecasts of future wages are based on these current wages, although these forecasts need not be optimal. Aloysius Siow (1984) and Gary Zarkin (1985) have extended Freeman’s framework by imposing the restriction that the expectations of future earnings are rational forecasts.

This paper attempts to synthesize the logit utility maximizing framework and the maximum expected earnings framework. We thus attempt to develop an empirically tractable model of occupational choice that more completely characterizes the underlying theoretical model. In addition, we consider the concurrent decision of whether to invest in further education. An agent decides which activity to select based on his expected utility from each career option, given tastes for future consumption, the existing stock of human capital, the hedonic attributes of the activity, the nonwage assets, attitudes toward risk, and the distribution of earnings within each activity. Future earnings are not known with certainty because the agent does not know where he will end up in the distribution of earnings. By specifying the form of the utility function, we can derive estimable equations relating the agent’s probability of choosing activity i to the moments of the distribution of earnings in activity i and to the agent’s past accumulation of human capital. This formulation may be used to justify jointly estimable reduced form occupational and educational choice equations. In addition, by imposing appropriate restrictions on the parameters of the model, structural parameters are derived which measure the level of embodied human capital in each activity and the supply responses to changes in the moments of the earnings distributions.

The model is tested using pooled time series and cross-section data which summarizes the occupational and educational choices of high school graduates in Maryland from 1951 through 1969. In general, the model generates reasonable supply elasticities. The results are consistent with the view that career choice is subject to risk averse behavior. Our results also support the hypothesis that schooling has activity-specific as well as general training components.

The next section introduces the theoretical framework. The third Section discusses the derivation of the earnings distribution per unit of human capital. Section four discusses the data, and Section five presents the results.
2. THE MODEL

We begin by assuming that individuals live for three periods. In the first period, individuals specialize in the accumulation of human capital. This human capital may be applied toward various occupational or educational opportunities. This human capital need not be equal across these activities since schooling may not be equally productive in all activities. Therefore, if there are $n$ potential activities an individual will finish period one with potential human capital stocks $h_i$, $i = 1, 2, \ldots, n$.

In period 2, individuals may invest in additional occupation-specific human capital. Thus, if $h_i$ represents the stock of occupation $i$-specific human capital, then an individual can generate a career level of occupation $i$-specific human capital equal to $H_i = (1 + g_i)h_i$. The coefficient $g_i$ measures the expected rate of growth of human capital while in occupation $i$. It represents the expected firm investment in occupation-specific human capital necessary to adequately perform on the job. Similarly, there is an associated expected growth rate in human capital from investing in vocational schooling or a university education during period two. If firms invest nothing in human capital of type $i$, $g_i = 0$ and $H_i = h_i$. The coefficient can be negative if $h_i$ more than meets the requirements for the job. In that case, human capital is allowed to depreciate, e.g., the firm invests in human capital at below the level required to maintain $h_i$.

The individual enters period three with a stock of human capital, $H_i$. Labor earnings in period three will equal $W_iH_i$ where $W_i$ is the discounted lifetime earnings per unit of human capital associated with activity $i$. If the activity is further formal education, say college, the relevant life-time earnings are the discounted earnings of college graduates. This allows symmetric treatment of the choice involving immediate entry from high school to an occupation and entry into a two- or four-year college.

In the empirical work that follows, we observe individuals at the end of period one as they are deciding which activity to undertake. The individual’s choice involves selecting an occupation or education track $i$ so as to maximize expected lifetime utility. We specify the choice as:

$$
\text{max } \{U_i(Y_i^* + (1 + g_i)h_i^*W_i)\}
$$

given $Y_i^*$, the savings from endowment income in period one, $h_i^*$, the human capital produced in period one, and also given knowledge of the $g_i$ and the distributions of $W_i$ for all $i$.

To operationalize (1) we specify the utility function as

$$
U_i = \alpha_i - \beta \exp (-\gamma H_i W_i - Y_i^*)
$$

3 The case of $g_i = 0$ implies that period two lasts no time and we enter immediately into period three. We should emphasize at this point that we assume implicitly that the intensity of investment in human capital in period two is determined by the choice of activity, e.g. the choice of entering a craft occupation is tied to a mandatory intensity of investment in craft skills in order to perform the job. Alternatively, one could make intensity of investment and the selection of activity joint choices, although this would complicate the analysis and subsequent empirical work considerably.
where $\alpha_i$ is a parameter representing the positive or negative hedonic return to activity $i$, $\beta$ is a taste parameter for period 3 consumption, and $\gamma$ is a measure of absolute risk aversion.\footnote{See Arrow (1971) for a general discussion of the theory of risk. Deaton and Muellbauer (1980) discuss the use of exponential utility functions in deriving demand systems for risky assets. Saito (1977) provides an empirical study of asset demand utilizing this framework.} Both $\beta$ and $\gamma$ are assumed to be positive.

The individual’s earnings from activity $i$ will not be known with certainty because of unknown future labor demand and supply conditions and the individual’s unknown relative future success in the activity. Even if earnings rates were known, future hours of work would not be known with certainty. In general, therefore, activity choice will depend on tastes for the activity, the availability of job training, and the distribution of earnings in the activity. Choices will depend only on the individual’s relative tastes for the activity if $\beta$ is zero, meaning the individual gets no utility from period 3 consumption, or if $Y^*_i$ approaches $+\infty$. Equation (2) is a concave, monotonically increasing function in wages. The function approaches $\alpha_i$ asymptotically. As $Y^*_i$ increases, the function shifts leftward and upward, meaning that $U_i$ approaches $\alpha_i$ more rapidly. In the limit, as $Y^*_i \to \infty$, the function will be a horizontal line at $\alpha_i$. Therefore, we would expect that taste will dictate the choice of activity to the greatest extent for the wealthiest segments of the population, while relative earnings distributions will dictate the choice for poorer segments of the population.\footnote{A similar argument is advanced by Freeman (1975a).}

We assume that the individual knows $Y^*_i$, $g_i$, and $\alpha_i$ so that these are fixed parameters in the individual’s decision process. We can write the expectation of $U_i$ as

$$E(U_i) = \alpha_i + E[-\beta_1 \exp (-\gamma H^*_i W_i)]$$

where $\beta_1 = \beta \exp (-Y^*_i)$. Taking the MacLaurin series expansion about $-\gamma H^*_i$ in the stochastic part of (3),\footnote{Because utility is ordinal rather than cardinal, we can expand the Taylor series about any point, $-\gamma H_i$. If $E(U_i) > E(U_j)$, $i \neq j$, evaluated at a point $-\gamma H_i$, then $E(U_i) > E(U_j)$ evaluated at any point $(-\gamma H_i) \neq (-\gamma H_j)$. We are implicitly assuming that whatever the distribution of the wage per unit of human capital in occupation $i$, the distribution is such that it has a moment generating function.} and assuming the distribution of $W_i$ has at least $m$ moments, we may approximate the expected utility of selecting activity $i$ as

$$E(U_i) = \alpha_i - \beta_1 \left[ 1 + E(W_i)(-\gamma H^*_i) + \frac{E(W_i^2)}{2} (-\gamma H^*_i)^2 + \cdots 
+ \frac{E(W_i^m)}{m!} (-\gamma H^*_i)^m + \cdots + \xi_i \right]$$

$$= \alpha_i - \beta_1 - \beta_1 \sum_{k=1}^{m} \frac{(-\gamma H^*_i)^k}{k!} E(W_i^k) + \beta_1 \xi_i$$
where $\xi_i$ is an error term representing the difference between our approximation and the actual expected utility. Notice that in this formulation, the mean of the distribution of earnings per unit of human capital serves to increase expected utility of selecting activity $i$. The second moment of the distribution decreases expected utility. A positive (negative) third moment of the distribution increases (decreases) expected utility, and so on. Notice also that we have placed no restrictions on the distribution of earnings other than that the moments of this distribution exist. In this formulation, we do not have to specify the distribution of wages explicitly.

The agent wants to pick the activity which will maximize his expected utility in period 3. Thus, the probability of choosing track $i$ over track $j$ may be written $P_i = Pr(E(U_i) > E(U_j))$. Since $\beta_1$ is deterministic and equal across all activities, this can be rewritten as

$$P_i = Pr \left[ \frac{E(U_i)}{\beta_1} > \frac{E(U_j)}{\beta_1} \right],$$

or

$$P_i = Pr \left[ (-\xi_i + \xi_j) < \frac{(\alpha_i - \alpha_j)}{\beta_1} - \sum_{k=1}^{m} \left( \frac{(-\gamma H^*_i)^k}{k!} E(W^*_i) - \frac{(-\gamma H^*_j)^k}{k!} E(W^*_j) \right) \right].$$

If we also assume that the activity-specific approximation errors $\xi_i$ and $\xi_j$ have independent Weibull distributions, $^8$ (5) may be written

$$P_i = \frac{\exp E(U_i) - \xi_i}{\sum_{j=1}^{n} \exp E(U_j) - \xi_j},$$

where $n$ is the number of available activities.

We can thus write a system of estimable equations for activities 1, 2, ..., $n - 1$, such that

$$\text{log} \left( \frac{P_i}{P_n} \right) = -\sum_{k=1}^{m} \left[ \frac{(-\gamma H^*_i)^k}{k!} E(W^*_i) - \frac{(-\gamma H^*_n)^k}{k!} E(W^*_n) \right] + \frac{(\alpha_i - \alpha_n)}{\beta_1}$$

for $i = 1, 2, ..., n - 1$. Equation (7a) shows that the log of the odds of choosing activity $i$ over activity $n$ is a function of the differences between the first through $m^{th}$ moments of the earnings distributions for activities $i$ and $n$ and the difference in taste between the activities $i$ and $n$. $^9$

$^7$ Although our economic agent may observe all moments of the distribution, in practice, the econometrician may only observe some finite number of these moments. Thus, the econometrician can only approximate the agent’s expected utility with error. We assume that this approximation error is random across time and space.

$^8$ For a description of the properties of the Weibull distribution, see Domencich and McFadden (1975).

$^9$ We do not formally treat the question of switching activities because we are concerned with explaining initial career choices. However, the initial career choice does not preclude subsequent changes
While individuals know the taste parameters $\alpha_i$ and $\alpha_n$, they are not known by the econometrician. To obtain an estimable version of (7a), we assume that the utility obtained by selecting activity $i$ over activity $n$ varies over time. We decompose the taste parameters into two components so that $1/\beta_1(\alpha_i - \alpha_n) = \bar{\varepsilon}_i + \varepsilon_i$, where $\bar{\varepsilon}_i$ is the mean preference for activity $i$ over activity $n$, and $\varepsilon_i$ is some time-specific positive or negative increment in the taste for activity $i$ over activity $n$. We can thus rewrite (7) as

$$
(7b) \quad \log \left( \frac{P_i}{P_n} \right) = - \sum_{k=1}^{m} \left[ \frac{-\gamma H_f^k}{k!} E(W_i^k) - \frac{-\gamma H_n^k}{k!} E(W_n^k) \right] + \bar{\varepsilon}_i + \varepsilon_i,
$$

$$
\quad \quad i = 1, \ldots, n - 1.
$$

If $H_f^k$ and the moments of the distribution of earnings per unit of human capital may be observed empirically, the structural parameter $\gamma$ may be identified from the restrictions implied by (7b). In general, the stock of human capital is not directly observable. However, inputs into the production of human capital (school quality, school attendance, family inputs) are observable.

We use a very simple formulation for our human capital production process, namely

$$
(8) \quad h_t = \delta_{0i} + \delta_{1i} \log (I_1 \cdot S_1).
$$

$I_1$ is a measure of time investment in schooling during period one and $S_1$ is a measure of school quality. We did experiment with functions involving additional inputs and different functional forms but found that the results were not overly sensitive to changes in specification. Provided $h_t$ is linear in parameters, the production function may contain any number of higher order polynomial terms so that the production function can be quite general.$^{10}$

Inserting (8) into (7b), and using the definition for $H_i$, we have

$$
(9) \quad \log \left( \frac{P_i}{P_n} \right) = - \sum_{k=1}^{m} \frac{(-\gamma)^k}{k!} \left[ \delta_{0i} + \delta_{1i} \log (I_1 \cdot S_1)(1 + g_i) \right]^k E(W_i^k)
$$

$$
\quad \quad - \left[ \delta_{0n} + \delta_{1n} \log (I_1 \cdot S_1)(1 + g_n) \right]^k E(W_n^k) + \bar{\varepsilon}_i + \varepsilon_i,
$$

$$
\quad \quad i = 1, \ldots, n - 1.
$$

If we ignore the activity-specific taste parameters $\bar{\varepsilon}_i$, the system of $n - 1$ equations has $3n + 1$ structural parameters including the absolute risk aversion parameter, $\gamma$, and the activity-specific human capital production parameters $\delta_{0i}$, $\delta_{1i}$, and $g_i$. Because of nonlinear dependence among the restrictions within and

---

$^{10}$ For example, $h_t$ can be specified as $\log (h_t)$ or $\exp (h_t)$, so long as the measure of human capital is linear in the parameters of the production function.
across equations, only $2n$ of the parameters are identified. Our strategy was, first, to fix $\gamma$ since it must be positive if individuals are indeed risk averse and since it serves mainly to scale the remaining parameters up or down by the same factor. Second, since $H_i$ and $h_i$ must be positive, $(1 + g_i)$ acts as a positive scaling variable on $\delta_{ii}$ and $\delta_{ij}$. We therefore estimate the $2n$ parameters, $(1 + g_i)\delta_{ii}$ and $(1 + g_i)\delta_{ij}$. These parameters are interpretable as the impact of period one human capital inputs on life-time activity-specific human capital, $H_i$. The number of parameters in the unrestricted form of (9) increases in the $m$ number of moments. For example, there are $4(n - 1)$ estimable parameters if $m = 1$; there are $10(n - 1)$ parameters if $m = 2$; there are $18(n - 1)$ parameters if $m = 3$, and so on. Thus, the structural parameters are just identified if $n = 2$ and $m = 1$. Increasing either $n$ (the number of occupations) beyond two or $m$ (the number of moments of the earnings distribution) beyond one will lead to overidentification of the structural system.

Because we can estimate the parameters of the underlying human capital production function, we can test whether school quality has different impacts on human capital across occupations. In other words, we can test whether general training has activity-specific components. We can also test whether school characteristics influence the mix of activities selected by the population. Finally, as a check on the theory, we can determine if our estimates of embodied activity-specific human capital $H^*_i$ are indeed positive and if the response of career choice to the moments of the distribution of earnings is in accord with the theoretical responses.

Although it would be preferable to include all the moments of the distribution of earnings per unit of human capital, the successive inclusion of higher order moments may eventually result in multicollinearity problems. Even with simple forms of the human capital production function of the type imposed herein, we were unable to admit more than two moments of the distribution. When the third moment was included, the moment matrix of the regressors became singular. For this reason, we restrict $m$ to be two in the estimation reported below.\textsuperscript{11}

3. ESTIMATING THE DISTRIBUTION OF EARNINGS PER UNIT OF HUMAN CAPITAL

The relevant distribution of earnings within an activity will undoubtedly differ across individuals because of differences in human capital accumulation in periods one and two. Unfortunately, we do not observe different distributions of earnings for individuals. We only observe the distribution of earnings for an entire occupation or educational group. It is important, therefore, to derive individual wage distributions from available information on the national market for labor within the activity and the characteristics of an individual’s human capital investment.

We begin by assuming that the distribution of earnings per unit of human capital within an activity is independent of the distribution of human capital within the activity. This assumption is equivalent to one in which the return from a given share

\textsuperscript{11} These problems are similar to those commonly encountered in empirical studies of production functions utilizing flexible functional forms.
of stock is independent of the distribution of stock holdings among shareholders. Mathematically, our independence assumption implies

$$f(W, H) = f_1(W)f_2(H),$$

where we have dropped the subscript $i$ for ease of exposition. Equation (10) shows that the joint density function of human capital and earnings per unit of human capital in an activity is equal to the product of the marginal densities of $W$ and $H$. This property implies that for an individual with human capital stock $H$, the conditional distribution of wages per unit of human capital is

$$f(W|H) = \frac{f_1(W)f_2(H)}{f_2(H)} = f_1(W)$$

so that the distribution of earnings per unit of human capital for any individual is the same, regardless of the individual’s holding of human capital. This property allows us to characterize the individual’s period 3 earnings as $H_2W$, which is identical to the formulation imposed in the previous section.

The assumption of independent density functions for $W$ and $H$ also allows a convenient transformation of available data to arrive at estimates of the moments of the distribution for $W$. We observe the distribution of $HW$ in an occupation. Provided the distributions of $W$ and $H$ have moment generating functions, the joint density of $W$ and $H$ is $M(W, H) = M(W, 0)M(0, H)$. If we restrict our analysis to the first two moments of the distribution, we can write

$$\mu_W = \frac{\mu_{W,H}}{\mu_H}$$

and

$$\sigma_W = \frac{\sigma_{W,H}}{\sigma_H}$$

where $\mu_j$ represents the first moment and $\sigma_j$ represents the second moment of the variable $j$. Thus, if we can obtain the distribution of human capital within an occupation as well as the joint distribution of human capital and earnings per unit of human capital in the activity, we may derive the moments of the distribution function for $W$. We first take logarithms of equations (12a) and (12b) to get equations

$$\ln \mu_{W,H} = \ln \mu_H + \ln \mu_W$$

and

$$\ln \sigma_{W,H} = \ln \sigma_H + \ln \sigma_W.$$
human capital on the logarithm of the respective moment of the distribution of human capital. While human capital in an activity is not directly observed, its determinants are observed. We can therefore obtain our estimate of $\mu_W$ and $\sigma_W$ from regressions of the mean and second moment of earnings in an activity on the means and second moments of the determinants of human capital in the activity. The procedure is outlined in the Appendix.

Given suitable measures of $\mu_W$ and $\sigma_W$, we must now attempt to construct an agent’s perception of the lifetime path of the distribution of wages in the occupation. Thus far, we have treated period three as a single year. In actuality, period three should be composed of $T$ subperiods, where $T$ is the length of time the individual expects to remain in the workforce.

Our strategy was to approximate the discounted lifetime earnings moments from activity $i$ as $(1 + \rho + \rho^2 + \cdots + \rho^T)M_i$ where $\rho$ is the discount factor. This can be approximated by $(1/(1 - \rho))M_i$ as $T$ becomes large. For junior college graduates, we subtract out the first two years of discounted earnings, so the measure becomes $(\rho^2/(1 - \rho))M_i$. Similarly, for four-year degrees, the sum becomes $(\rho^4/(1 - \rho))M_i$.

In an earlier version of this paper, we generated discounted earnings paths based on expected future movements in the moments of the earnings distributions. In general, these rational projections yielded similar results for occupational earnings distributions but did not yield reasonable figures for the earnings distributions for the college and junior college options. In addition, the simpler estimates proved more successful in explaining career choices than did the rational projections. We therefore restrict our discussion herein to results generated using the simpler discounted sums.\(^{12}\)

4. DATA

The model is estimated over a sample of 23 school districts in Maryland from 1951 through 1969. The Maryland State Board of Education each year reported the proportion of recent male high school graduates in each of the 23 counties who were engaged in each of eight activities. These included graduates who took jobs in six broad occupations: farming, fishing and mining; operatives and laborers; service workers; crafts; clerical and sales; and professionals and managers. In addition the schools reported two academic groups: the proportion attending vocational and technical schools; and the proportion attending four-year colleges and universities. The data are particularly well suited for our study since the time span is sufficiently long to allow for substantial changes within and between the occupational wage distributions. The cross-section allows us to control for location-specific, time-invariant effects as well as differences in family and school inputs into human capital production. Means and standard deviations of the variables are reported in Table 1.

The dependent variable in (9) is measured as the logarithm of the proportion of

\(^{12}\) Results using optimal projections of the Hansen and Sargent (1980) type are available from the authors on request. Siow (1984) and Zarkin (1985) have examined rational expectations models in the supply of lawyers and teachers respectively.


<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log ((P_i/P_0)^*), (P_i =)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional</td>
<td>2.31</td>
<td>2.01</td>
</tr>
<tr>
<td>Operatives/Laborers</td>
<td>.77</td>
<td>1.29</td>
</tr>
<tr>
<td>Clerical/Sales</td>
<td>.09</td>
<td>1.39</td>
</tr>
<tr>
<td>Craft</td>
<td>-.67</td>
<td>1.58</td>
</tr>
<tr>
<td>Service</td>
<td>-1.56</td>
<td>1.80</td>
</tr>
<tr>
<td>Junior College/Vocational</td>
<td>-.72</td>
<td>1.57</td>
</tr>
<tr>
<td>University</td>
<td>1.43</td>
<td>1.42</td>
</tr>
</tbody>
</table>

**Independent Variables:**

<table>
<thead>
<tr>
<th>Log of Instructional Salaries**</th>
<th>5.50</th>
<th>.26</th>
</tr>
</thead>
</table>

**First Moment of Income Distribution**

<table>
<thead>
<tr>
<th>Professional</th>
<th>1.01</th>
<th>.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operative/Laborers</td>
<td>.65</td>
<td>.07</td>
</tr>
<tr>
<td>Clerical/Sales</td>
<td>.77</td>
<td>.09</td>
</tr>
<tr>
<td>Craft</td>
<td>.77</td>
<td>.08</td>
</tr>
<tr>
<td>Service</td>
<td>.53</td>
<td>.06</td>
</tr>
<tr>
<td>Junior College/Vocational</td>
<td>.73</td>
<td>.16</td>
</tr>
<tr>
<td>University</td>
<td>.92</td>
<td>.17</td>
</tr>
<tr>
<td>Farming, Fishing &amp; Mining</td>
<td>.55</td>
<td>.12</td>
</tr>
</tbody>
</table>

**Second Moment of Income Distribution**

<table>
<thead>
<tr>
<th>Professional</th>
<th>.98</th>
<th>.13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operative/Laborers</td>
<td>.33</td>
<td>.07</td>
</tr>
<tr>
<td>Clerical/Sales</td>
<td>.51</td>
<td>.11</td>
</tr>
<tr>
<td>Craft</td>
<td>.45</td>
<td>.09</td>
</tr>
<tr>
<td>Service</td>
<td>.25</td>
<td>.05</td>
</tr>
<tr>
<td>Junior College/Vocational</td>
<td>.46</td>
<td>.14</td>
</tr>
<tr>
<td>University</td>
<td>.77</td>
<td>.24</td>
</tr>
<tr>
<td>Farming, Fishing &amp; Mining</td>
<td>.38</td>
<td>.18</td>
</tr>
</tbody>
</table>

\(^*P_n\) = Proportion in Farming, Fishing and Mining

\(^**\) per attending student multiplied by the average attendance rate (six-year moving average)

students selecting activity \(i\), relative to the proportion in farming, fishing and mining. Two observations were dropped due to apparent reporting errors, leaving 435 observations for each of the seven activity choice equations.

The Maryland State Board of Education also reported data on school quality and attendance for each county. For our measure of school inputs into human capital production, we use instructional salaries per student, denominated in 1967 dollars. To control for differences in time spent in human capital investment across counties, we multiplied this measure of school quality by the average annual attendance rate in each county. Finally, to account for the fact that human capital production in schools took place over several years, we took a six-year moving average of the school input measure as our measure of \(I_1\cdot S_1\) in equation (9).13

Data on the yearly earnings distribution for each occupation were taken from the

13 The use of the six-year average was arbitrary, but the measures of school quality and household inputs should not be overly sensitive to changes in the specification of the moving averages.
U.S. Bureau of the Census, Series P-60. The income distributions were reported by discrete intervals from which we obtained measures of the moments of the distribution using standard procedures. Data on the distribution of educational attainment within each occupation were obtained from the Bureau of Labor Statistics. The mean and second moments of the income distributions, valued in constant dollars and purged of the influence of shifts in the distribution of educational attainment and age, are our measure of the moments of the distribution of earnings per unit of human capital, \( W_j \), in equation (9).

To control for differences in tastes both across activities and across counties within activities, we include a separate fixed-effect for each county in each career choice equation. Since estimation of these effects becomes unduly expensive in maximum likelihood estimation (they add 161 parameters to the model), we control for these fixed effects by taking deviations from the county-specific means of all the variables used in estimation. Thus, our structural parameter estimates are based on data in this differenced form. Alternatively, we could have used a random-effects model of the type advanced by Balestra and Nerlove (1966) and Maddala (1971). A recent paper by Jakubson (1988) found that fixed-effect and random-effect life-cycle labor supply estimates yielded similar results although excluding controls for individual heterogeneity led to significant biases in the parameter estimates.

5. STRUCTURAL ESTIMATION OF THE CAREER CHOICE MODEL

In this section, we report the results of our empirical investigation of the career decisions of male high school graduates. The estimates are generally consistent with expectations. The results indicate that higher order moments of the earnings distribution do enter the career choice decisions, as does the level of human capital. Furthermore, the human capital produced from a given level of school inputs varies across activities. In other words, the training obtained in school or in the household has activity-specific effects as well as general effects.

The equation which we estimate is a special case of equation (9) in which \( k \) is set equal to 2. The moments are generated as described in equations (13a) and (13b) and the Appendix. We assume that the \( e_{ij} \) have a multivariate normal distribution with uncorrelated disturbances across time and space. The covariance matrix is specified as \( E(e_{ij} e'_{ij}) = V \).

The concentrated likelihood function over \( T \) time periods, \( c \) counties and \( n - 1 \) equations is

\[
L(\theta) = -((n - 1)mTc/2)(1 + \log (2\pi) - \log (Tc)) - (Tc/2) \log \left| \sum_{i=1}^{c} M_i(\theta) \right|
\]

\[14\] Our strategy will yield consistent but not efficient estimates of the parameters. This means that the standard errors generated herein should be interpreted cautiously.
where $M_i(\theta)$ is the $(n - 1)$ by $(n - 1)$ moment matrix of residuals for the individual $i$ corresponding to the vector of parameters, $\theta$. Equation (14) may be maximized with respect to $\theta$ to obtain the maximum likelihood estimates of the structural parameters. As noted above, the value of $\gamma$ must be positive if agents are risk averse. We fixed its value at .25. Because $\gamma$ serves to scale the other parameters, the choice of $\gamma$ will not alter significance tests or elasticity estimates. For example, raising $\gamma$ to .5 will cut in half both the human capital production process parameter estimates and their associated standard errors.

Once $\gamma$ is fixed, our least restrictive parameter vector is $\theta_1 = [\delta_{0i}(1 + g_i), \delta_{1i}(1 + g_i)]'$, $i = 1, 2, \ldots, 8$. To test if school quality has activity-specific effects on human capital, we consider a subset of $\theta_1$, $\theta_2 = [\delta_{0i}(1 + g_i), \delta_{1i}(1 + g_i)]'$ which restricts that school inputs have the same productivity across all activities. Letting $L(\cdot)$ be the likelihood value for the relevant parameter vector, the test statistic for the hypothesis of neutral effects of school inputs on choice of activity is $2(L(\theta_1) - L(\theta_2))$. The test statistic is distributed $\chi^2(7)$, where 7 is the number of restrictions.

The model was first estimated without imposing the within-equation and cross-equation restrictions. Some mild evidence of first-order autocorrelation in the errors was observed, so the estimated model was transformed by multiplying both sides of equation (9) by $(1 - \rho_i) L$ where $L$ is the lag operator and $\rho_i$ is the coefficient in the first-order process generating the error structure, $\epsilon_{ijt} = \rho_i \epsilon_{ijt-1} + \eta_{ijt}$. Not including the fixed-effects, the unrestricted form of the model had 21 parameters in each equation and 147 parameters over seven equations. This form of the model was estimated using Zellner’s seemingly unrelated equations technique. The model performed quite well, with 59 to 87 percent of the variation in each choice equation explained by the regressors. The results clearly demonstrate the need to consider more than just the first moment of the earnings distribution in analyzing career choice. The null hypothesis that the second moment of the distribution of our computed earnings per unit of human capital did not enter the system of career choice equations was strongly rejected. The results also demonstrated that variation in school inputs into human capital production during the period also significantly influenced occupational and educational choices, holding constant the years of schooling and county-level fixed effects. Therefore, studies of occupational choice that control for human capital production only by holding constant the years of schooling fail to adequately capture the role of human capital production in influencing occupational choices.

The unrestricted estimates are most useful for forecasting purposes or for joint tests of the type reported above. The parameters themselves are difficult to interpret due to the large number of interaction terms. The structural restrictions offer a means by which explicit interpretations may be placed on the model’s parameters. The validity of these restrictions can be tested by a standard likelihood ratio test.

The maximum likelihood estimation was performed using the DFP algorithm.

---

15 The derivation of the concentrated likelihood function follows the procedure outlined in Bard (1974).

16 These unconstrained estimates are not reported but may be obtained from the authors on request. Estimates which constrain only $\delta_{ij} = \delta_{in}$ in equation (9) are reported in Orazem and Mattila (1986).
from the GQOPT2 program developed at Princeton University. The results from the least restrictive parameter set \( \theta_1 \) are contained in Table 2. There are no strong priors on the signs of the constant terms, but the coefficients on the schooling input should be positive. However, two of the school production coefficients, those in the operative occupation and in the craft occupation, are negative and significant. The literal interpretation of the findings is that secondary schooling reduces occupation-specific human capital in these occupations, and that school inputs are most important in generating human capital in the junior college, service, clerical and university activities. The null hypotheses that school quality had the same effect across all activities was easily rejected. The implication of this finding is that schooling has activity-specific as well as general training components.

Given the sparse specification of the production process, the results in Table 2 may be more plausibly interpreted as parameters in a first-order approximation to an unknown human capital production function evaluated at a specific point.
Therefore their main value lies not in their individual interpretations but in their combined implications regarding activity-specific levels of human capital and in revealing how choice of educational or occupational track respond to the means and variances of earnings distributions.

The level of human capital production in each occupation may be estimated by evaluating the equation, $\delta_{0i} + \delta_{1i} \log (I_i \cdot S_i)$, at the sample means reported in Table 1. We find that our estimates of human capital production are positive in all eight educational and occupational tracks. The estimates are reported in Table 3.\textsuperscript{17}

Since each measure of human capital is specific to its activity, the measures are not comparable across activities. We therefore multiply the measures of human capital by the corresponding average earnings per unit of human capital. The results are listed in the second column of Table 3. The results indicate roughly similar values of human capital from entering the professional, operative, craft, and farming occupations, somewhat lower value of applying human capital to the clerical occupation, and the lowest value to applying the human capital toward the service occupation and the two educational tracks.

The finding of smaller returns to human capital in the two educational activities merits additional comment. The explanation from the model is that the hedonic attributes from furthering one’s education must dominate the hedonic attributes of entering an occupation immediately upon graduation from high school. Therefore, there is an overinvestment in education on purely pecuniary grounds. This is, in fact, an argument advanced by Richard Freeman (1976). Those who attended college in this period may have had the highest stocks of human capital on average.

\textsuperscript{17} As an example of these estimates, the human capital estimate for professionals is $\delta_{0}$ professional + $\delta_{1}$ professional * log ($I_i \cdot S_i$) = 31.48 - 4.6 * 5.5 = 6.18.
but, according to the results in Table 3, would have even higher returns on their human capital had they selected an occupation straight out of high school.

The supply elasticities for each occupation are also reported in Table 3. In all eight cases, we find that increasing the first moment of the \( i \)th earnings distribution increases the probability of selecting the \( i \)th activity and increasing the second moment of the \( i \)th earnings distribution reduces the probability of selecting the \( i \)th activity. The supply response is generally elastic in the occupations and inelastic in furthering general training through vocational or college education. Educational choices are also less sensitive to changes in second moments of the distribution of earnings than are occupational choices.

We also computed how the log odds of selecting the \( i \)th activity relative to farming varied with changes in school quality. The elasticity was computed at sample means, taking the derivative of (9) with respect to \( \log (I_1 \cdot S_1) \), as \( \gamma (\delta_{11} \mu_{wi} - \delta_{1F} \mu_{ni}) - \gamma^2 (\delta_{11} H_{1i} \sigma_{wi} - \delta_{1F} H_{1i} \sigma_{ni}) \). Here we have the opposite pattern in the relative size of the responses of occupational and educational choices. Educational choices are more sensitive to secondary school quality and attendance. In general, raising school inputs tends to increase the proportion of high school graduates entering all choices relative to farming, fishing and mining.\(^{18}\)

To test the validity of the structural restrictions imposed by the model, we compare the likelihood values in Table 2 with the likelihood value of the unrestricted estimation. The restrictions are rejected.\(^{19}\) It should be stressed that this test rejects the simple human capital production function we imposed in equation (9), since all of the within equation and cross-equation restrictions result from our assumptions concerning the form of the production function. The theory itself is not rejected, and in fact, the unrestricted form of (9) may be viewed as a more general approximation to the underlying theoretical activity selection equation represented

\(^{18}\) Our finding of an elastic supply response to changes in the mean earnings in the professional occupation is consistent with findings reported by Freeman (1975a, 1975b) for lawyers and physicists. Estimates of entry-level supply responses for the other occupations are not readily available in the literature. Of course, some might wonder if the Maryland data are representative of the United States. The following comparison indicates that occupational distributions are quite similar. The Maryland figures are averages for high school graduates over the period 1951–70. The U.S. data are averages for males 18–19 having exactly four years of high school in 1960 (U.S. Bureau of Census, 1960 Census of Population, Educational Attainment, Report PC(2)-5B, Subject Reports, Table 8).

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Maryland</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professionals &amp; Managers</td>
<td>.020</td>
<td>.058</td>
</tr>
<tr>
<td>Clerical &amp; Sales</td>
<td>.215</td>
<td>.251</td>
</tr>
<tr>
<td>Craftsmen</td>
<td>.101</td>
<td>.120</td>
</tr>
<tr>
<td>Operatives &amp; Laborers</td>
<td>.426</td>
<td>.414</td>
</tr>
<tr>
<td>Service</td>
<td>.041</td>
<td>.076</td>
</tr>
<tr>
<td>Farm &amp; Farm Laborers</td>
<td>.197</td>
<td>.081</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

The discrepancy for “farmers” is due partly to the fact that the classification scheme for Maryland includes mining, fishing, and lumbering in addition to farming, whereas the U.S. classification is exclusively farmers and farm laborers.

\(^{19}\) The log-likelihood value for the unrestricted model is \(-3522.6\). Twice the difference of the likelihood values between the unrestricted and restricted models is 246 which exceeds the critical values of the test at the .05 level of confidence.
by (7). Still, the structural estimates reported here seemed quite reasonable. These results would indicate that the restrictions imposed by more general functional forms may be supported by the data. Further experimentation with more general forms of production functions would appear to be warranted.

6. CONCLUDING COMMENTS

This paper develops an empirical model of occupational and educational choice under uncertainty. The model yields a multinomial logit specification of the career choice decision that is directly derived from the underlying theory. This specification shows that the decision will depend on the distribution of earnings per unit of human capital and not just the path of the mean of the distribution of earnings in the activity. In addition, it was demonstrated that changes in the level of human capital production may also alter an agent’s career choice decision.

The results clearly show that higher-order moments of the distribution of earnings in an occupation do enter the career choice decision in the manner predicted by the theory. Furthermore, school quality has activity-specific effects, implying that changes in the level of human capital production will alter the supply of new entrants to a given activity. While the specification of the human capital production function was not supported by the data, the structural results yield reasonable estimates of the level of human capital production in each activity and supply elasticities with respect to the moments of the distribution of earnings per unit of human capital.

The results herein would support further experimentation with this type of model in studying human capital investment in the presence of uncertain returns. Extension of the model into a multiple period framework to allow for switching behavior would be one valuable path of research. Experimentation with alternative forms of the human capital production function may yield more satisfactory individual production parameters. Finally, the model can be utilized in its unrestricted form as a justification for a more complete specification of empirical models of occupational choice.

Iowa State University, U.S.A.

APPENDIX

Data on the earnings distribution by occupation are available from the U.S. Bureau of the Census, Series P-60. Income distributions by discrete intervals have been published annually for males by major occupation. From these distributions, we estimated the annual mean and the second moment of income (each expressed in real terms) for each of our six occupational groupings. These moments represent our measures of \( \mu_{W,H} \) and \( \sigma_{W,H} \).

Data on the distribution of educational attainment by occupation is available from the U.S. Bureau of Labor Statistics for 1952, 1957, 1959, 1962, and 1964–1970 on a consistent basis for males 18 years of age and older. Moments of the distribution for missing years were estimated by interpolation, which should be
reasonable since the distributions change slowly. The distribution of educational attainment is our proxy for the distribution of human capital in the occupation.

For each occupation, we ran ordinary least squares regressions of the form

\[ \ln \mu_{W,H} = a_1 + b_1 \ln \mu_s + \varepsilon_1 \]

\[ \ln \sigma_{W,H} = a_2 + b_2 \ln \sigma_s + \varepsilon_2 \]

where \( \mu_s \) and \( \sigma_s \) are the first and second moments of the distribution of schooling respectively, \( a_1 \) and \( a_2 \) are constants, and \( \varepsilon_1 \) and \( \varepsilon_2 \) are residuals. The constant term was included to control for time invariant differences in the distribution of human capital across occupations such as the distribution of ages within the
occupation. The error terms are, by construction, the portion of the first and second moment of the distribution of earnings that are uncorrelated with movements in the distribution of human capital. Therefore, using equations (13a) and (13b), \( \epsilon_1 = \ln \mu_W \) and \( \epsilon_2 = \ln \sigma_W \). We can thus obtain measures of the moments of the distribution of earnings per unit of human capital by taking the exponentials of \( \epsilon_1 \) and \( \epsilon_2 \).

A similar procedure was used to estimate the distribution of earnings per unit of human capital in the population of junior college and university graduates. For these groups, the logarithm of the first and second moments of earnings were regressed on a constant and the logarithm of the age distribution of the population of graduates.
Because the residuals will all have mean zero, the time series of exponentiated residuals will have a mean of one. While these paths give the trajectory of earnings per unit of human capital for each occupation over time, they do not capture the relative earnings across occupations at any one point in time since the first and second moment of earnings for each occupation will be one. Data from the 1950, 1960, and 1970 editions of the Census of Population were used to determine the relative values of the first and second moments of the distribution of earnings of high school graduates in each occupation. These relative values for 1949, 1959, and 1969 were used as benchmarks. The trajectory of earnings for each occupation in the intervening years was adjusted to meet these benchmark levels. The resulting time series of mean earnings per unit of human capital are shown in Figure A.1, and the time series of the second moments are shown in Figure A.2.

REFERENCES

MARYLAND STATE BOARD OF EDUCATION, Annual Report: A Statistical Review for the Year (Baltimore, annual issues).


