University of Wisconsin-Madison

IRP Discussion Papers

Christopher J. Flinn

WAGES AND JOB MOBILITY
OF YOUNG WORKERS

DP #728-83

Institute for Research on Poverty
Discussion Paper 728-83

Wages and Job Mobility of Young Workers*

Christopher J. Flinn

June 1983

*This research was partially supported by NSF Grant SES 81-07963 and a Sloan Foundation Grant to the Institute for Research on Poverty at the University of Wisconsin. The paper was prepared for the Hoover Conference on Labor Economics held in January 1983. I have received valuable comments from Gary Chamberlain, Jim Heckman, Glenn MacDonald, Bob Miller, Jacob Mincer, Eytan Sheshinski, and participants in seminars at the SSRC Workshop on Job Turnover and the University of Wisconsin-Madison. I am particularly grateful to Bob Shakotko for pointing out an error in an earlier draft. George Jakubsen and Dan Feaster provided valuable programming assistance and advice. Marie Goodman provided expert clerical assistance. I am responsible for any errors or inconsistencies that remain.

ABSTRACT

In this paper a model of wage determination and interfirm mobility decisions is exposited and estimated. The theoretical model is essentially a discrete time version of Jovanovic's worker-firm matching model. Workers only stay at firms in which the quality of the match is perceived to be high. Mobility decisions are made sequentially over the course of the employment spell as new information reflecting the quality of the match arrives.

Using a sample of young men from the National Longitudinal Survey, we jointly estimate a system of dynamic wage equations and parameters of the mobility decision rule. We find that worker-firm specific heterogeneity accounts for almost as much of the variability in log wages for recent market entrants as does individual specific, time and firm invariant heterogeneity. While differences in the quality of worker-firm matches are an important component of wage variability, the results suggest that matching heterogeneity cannot account for worker-firm separations occurring at tenure levels greater than two or three years.

Wage and Job Mobility of Young Workers

1. INTRODUCTION

It is a well-established empirical regularity that the probability of leaving a job is a decreasing function of a worker's tenure at a firm (see Jovanovic, 1978, Ch. 1, for a detailed review of empirical work on turnover). One explanation advanced is that the accumulation of firm-specific human capital makes a job change costlier for workers with more experience at a given firm (see, e.g., Becker [1962], Rosen [1968], Parsons [1972], and Kuratani [1973]). Another explanation for the tenure-turnover relationship is that there exists a productivity effect intrinsic to the match of a worker and a firm. While all agents are assumed to know the parameters of the distribution of this matching heterogeneity, the value of a particular match is only partially observable to both the worker and firm. The longer an employment spell continues, the more precise is the estimate of the value of the match. As a consequence of this learning process, turnover is more likely to occur at low tenure levels.

That worker-firm heterogeneity may be an important factor in the explanation of turnover was recognized by Silcock (1954). Jovanovic (1978) was the first to establish the form of equilibrium wage contracts and turnover decision rules for a matching model formulated in continuous time. Wilson (1980) has solved for equilibrium wage policies in a model that includes both matching heterogeneity and job search. Other nonequilibrium models that generate turnover are those of Wilde (1980), Johnson (1978), Burdett (1978), and Lippman and McCall (1978). In Burdett's model some employed individuals search and when superior offers arrive leave their current firm. Johnson's model of job shopping is essentially a two-period, two-firm, discrete time version of Jovanovic's model. The simplicity of his model allows him to obtain a number of comparative static results. In the models of Wilde and Lippman and McCall, workers learn a characteristic of the job only after accepting employment. If the value of this characteristic is sufficiently low, the worker quits.

The purpose of this research is to empirically assess the quantitative importance of worker-firm matching heterogeneity in explaining interfirm mobility and wage determination. The model of turnover and wage determination discussed and estimated in this paper is similar to those of Jovanovic (1978) and Johnson (1978). It is set in discrete time, as is the job-shopping model of Johnson, and assumes infinitely lived agents, as does Jovanovic.[1] In order to focus attention on the worker-firm matching process, we ignore firm-specific human capital accumulation in what follows. The introduction of firm-specific human capital greatly complicates the form of the turnover decision rule, and may produce a policy function which does not possess the reservation value property. For the theoretical model to provide any guidelines for our econometric specification, the turnover decision must possess the reservation value property.

A note on the organization of the paper. In Section 2 we state the nature of the worker's utility maximization problem, the content of the worker's information set, and establish that the worker's turnover deciworker's information set, and establish that the worker's turnover decision possesses the reservation value property. In Section 3 we discuss the data employed in the empirical analysis and present some descriptive statistics. Section 4 contains a discussion of the problem of obtaining initial consistent estimators for estimable structural parameters in the model. For the matching model presented here, some of the parameters are nonparametically underidentified, in the terminology of Flinn and Heckman (1982b). In Section 5 we obtain maximum likelihood estimates of the identifiable structural parameters after making specific distributional assumptions. Section 6 contains a brief conclusion.

Finally, a note on some of the inadequacies of the treatment of turnover in this paper. Because the focus of this work is primarily empirical, the approach taken here is a decidedly nonequilibrium one.[2] Firms are assumed to be passive agents in all that follows. The rents that accrue to the match are either entirely captured by the worker or divided between worker and firm in some constant proportion.[3] Only permanent separations are considered, and due to the passive firm assumption, all quits are voluntary.[4] Wages are assumed to be the only job attribute of value to the worker.

2. TURNOVER DECISIONS IN A WORKER-FIRM MATCHING MODEL

Workers are assumed to be infinitely lived, with entry into the labor market occurring at t=0 (a normalization). Consumers maximize the expected value of the discounted sum of a time separable utility function of the form

$$W_{it} = \max_{\{j_k\}_{k=t}^{\infty}} E_t \sum_{k=t}^{\infty} \beta^{k-t} U(c_{it})$$

$$\text{s.t. } c_{it} = w_{ijt},$$

where W_{it} is the value of the i^{th} agent's problem at time t, c_{it} is the consumption of agent i at time t, β is the time invariant discount factor, and w_{ijt} is the wage rate of agent i at firm j at time t. The conditional expectation operator E_t denotes expectation taken with respect to information available at the end of period t-1. Workers are constrained to consume all earnings during the period in which they are received.

At this point we make several functional form assumptions which considerably simplify the analysis. Assume that $U(c_{it}) = \ell n\ c_{it}$ and substitute the budget constraint into the objective function to get

$$W_{it} = \max_{\{j_k\}_{k=t}^{\infty}} E_t \sum_{k=t}^{\infty} \beta^{k-t} \ell n\ w_{ij_k k},$$

where the j subscript indexes firms in the agent's choice set, and assumes values j = 1,2, ..., J. We assume that all agents possess the same choice set, and that the set of firms is large but finite. The subscript j_k denotes the firm in which individual i is working at date k. The maximization procedure will be made explicit below.

The productivity of the i^{th} worker at the j^{th} firm at time t is given by

$$w_{ijt} = \exp\{\underset{\sim}{Z}_{it}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij} + \varepsilon_{ijt}\},$$

where $\underset{\sim}{Z}_{it}$ is a K × 1 vector of observable, individual specific, time-varying productivity characteristics, $\underset{\sim}{\gamma}$ is K × 1 parameter vector, η_i is an individual specific, time invariant heterogeneity component, θ_{ij} is a worker-firm productivity component, and ε_{ijt} is white noise. The values of η, θ, and ε are assumed to be unobservable to the analyst. We assume that all three unobservables η, θ, and ε possess absolutely continuous distribution functions; that they are independently distributed across workers, firms, and time; and that the parameters of the three distribution functions are not functions of observable characteristics of workers or firms.

The maximization procedure may now be described. Since individuals capture the rents which accrue to the worker-firm match, workers will attempt to locate the firm at which their value of θ_{ij} is largest. The constraints on worker "searching" are the following:

1) There may exist a direct cost of finding a new employer or terminating a current match. Denote this direct cost by $C(C \geqslant 0)$.

2) Once a worker leaves a firm he may not return to the firm (no recall). Less stringently, we may assume that a worker who returns to a previous employer draws a new match value (θ) independent of the previous value realized at the firm.

3) All firms look alike <u>ex ante</u> to workers. The value of the match of worker i to firm j provides no information as to the expected value of the match of worker i with firm j', $j' \neq j$. This has already implicitly been assumed above, but we restate it here since it has a direct bearing on the form of the worker's decision rule. Formally,

$$M_i(\theta_{i1}, \theta_{i2}, \ldots, \theta_{iJ}) = \underbrace{G_i(\theta_{i1}) \cdots G_i(\theta_{iJ})}_{J \text{ times}},$$

thus the joint distribution function of agent i's match values with any J firms is the product of J identical univariate distribution functions. As stated above, we assume that $G_i(\theta_{ij}) = G(\theta_{ij})$, i.e., all agents $i = 1, \ldots, I$ draw from the same match distribution. (Thus the match distribution is not a function of observable <u>or</u> unobservable characteristics.) Agents are assumed to perfectly observe the value of their time and match invariant productivity component, η_i, at the time of entry into the labor market. They can only observe the <u>sum</u> of the match value θ_{ij} and white noise ε_{ijt} each period. Define θ_{ijt} +

$\varepsilon_{ijt} \equiv r_{ijt}$. We shall establish the existence of a reservation value property for this model.

For a worker who had been working for his current employer j for d periods his date t turnover policy function will be of the form

(2.1) (S) stay with current firm iff $R(r_{ij,t-1}, \ldots, r_{j,t-d}, d) \geq R_d^*$.

(M) leave current firm iff $R(r_{ij,t-1}, \ldots, r_{ij,t-d}, d) < R_d^*$.

The decision rule only includes as arguments realizations of $\{r\}$ which occur at the current job, since match draws are i.i.d. and $\{\varepsilon\}$ is white.

Since all firms look alike <u>ex ante</u>, the choice the worker faces at each point t is between staying at the firm at which he was employed in period t-1 or moving to one of the remaining J-1 firms. Further, we assume that wages are bounded, which implies that all unobservables (to the analyst) η, θ, and ε have bounded support. Although we assume η, θ, and ε have bounded support, we assume each is <u>approximately</u> independently and normally distributed with mean zero for all (i, j, t).[5] The normality assumptions are essential to obtain tractable characterizations of the agent's information acquisition process. Most importantly for purposes of this section,

$$\theta_{ij} \sim N(0, \sigma_\theta^2) \ \forall \ i,j.$$

$$\varepsilon_{ijt} \sim N(0, \sigma_\varepsilon^2) \forall \ i,j,t.$$

Define the precision of θ and ε by Π_θ and Π_ε respectively, where $\Pi_k = 1/\sigma_k^2$, $k = \theta, \varepsilon$.

At the beginning of period t, the agent's decision to change employers will depend on his point estimate of the quality of his match with his current employer, q_{ijt}, and the precision of the point estimate, h_{ijt}. Due to the normality assumptions, the worker's posterior distribution of $\hat{\theta}_{ij}$ is characterized by these two quantities. For simplicity we drop the subscripts on q and h for the remainder of this section. Let q' be the updated point estimate of q given one additional observation of r on the current job; let h' similarly denote the updated precision of the sequentially revised point estimate q'. The relationships between these quantities are given by the well-known expressions (e.g., DeGroot, 1970):

$$(2.2) \quad q' = \frac{hq + \Pi_\varepsilon r}{h + \Pi_\varepsilon}$$

and

$$(2.3) \quad h' = h + \Pi_\varepsilon .$$

When a worker accepts employment with a new firm, the initial values of q and h are 0 and Π_θ respectively. The process of updating begins from these values for all new employers. One more observation is needed before deriving the turnover decision rule; given a current estimate of the match value q with precision h, the marginal distribution of the next observation r' is normal with mean q and variance $(h + \Pi_e)/ h\Pi_\varepsilon$.

Since the posterior distribution of the match value can be characterized by the mean and precision, the worker's mobility decision at each time t will only be a function of the current values of these two state variables. To investigate properties of the decision rule, first assume a finite horizon with terminal date T. We can write the T period value function as

(2.4) $V_T(q,h) = \max\{q;0\}$.

Obviously, V_T is increasing in q and is independent of the precision h. No matter what the value of h, the turnover decision rule is a simple one: stay with the current employer if $q \geq 0$; leave the current employer if $q < 0$. The reservation value at T is given by $q_T^*(h) = q_T^* = 0$.

The period T − 1 value function is

$$(2.5)\quad V_{T-1}(q,h) = \max\Big\{q + \beta\int V_T(q'(q,h,\phi),h')dS(\phi \mid q, \frac{h + \Pi_\varepsilon}{h\Pi_\varepsilon});$$

$$-C + \beta\int V_T(q'(0, \Pi_\theta, \phi), \Pi_\theta + \Pi_\varepsilon))dS(\phi \mid 0, \frac{\Pi_\theta + \Pi_\varepsilon}{\Pi_\theta\Pi_\varepsilon})\Big\},$$

where $S(\phi \mid a,b)$ denotes the cumulative distribution function of a normal random variable ϕ with mean a and variance b. As before, the first argument of the maximization operator in (2.5) is the value of staying at the current employer and the second is the value of taking a job with a new employer. The function $q'(x,y,z)$ is an implicit representation of the mean-updating equation (2.2) where x is the previous point estimate, y is the previous precision, and z is the new observation on r which is a random variable.

We may establish that V_{T-1} is increasing in q by observing the following. The second argument in the maximization operator is a constant independent of q, therefore we only need to establish that the first argument is increasing in q. This amounts to showing that

$$(2.6)\quad \int V_T(q'(q,h,\phi),\ h')dS(\phi \mid q,\ \frac{h + \Pi_\varepsilon}{h\Pi_\varepsilon})$$

is increasing in q. Since V_T is increasing in $q'(q,h,\phi)$ and $q'(q,h,\phi)$ is increasing in q, then $V_T(q'(q,\ h,\ \phi),\ h')$ is increasing in q for all ϕ. But increases in q also affect the distribution of ϕ. Consider $q_1 > q_2$. Then $S(\phi \mid q_1,\ b)$ first order stochastically dominates $S(\phi \mid q_2,\ b)$ for all b finite. This relationship implies $\int \ell(\phi)dS(\phi \mid q_1,\ b) \geq \int \ell(\phi)S(\phi \mid q_2,\ b)$ for all ℓ belonging to the class of increasing functions. Since V_T is a member of this class, the two effects of increasing q both act to increase (2.6). Thus we have shown that increases in q increase the first argument in the max operator in (2.5), which establishes V_{T-1} increasing in q.

Now consider the effects of increases in the precision h on V_{T-1}. Once again the second term in the max operator of (2.5) is independent of h, as is the expected current period return at the current employer. Increases in h only affect V_{T-1} through the term reproduced in (2.5). As before we first examine the effect of increases in h on EV_T, ignoring effects changes in h induce on the distribution of ϕ. Consider $h_1 > h_2$, and compare $\int V_T(q'(q,\ h_1,\ \phi),\ h_1 + \Pi_\varepsilon)dS(\phi \mid q,b)$ and $\int V_T(q'(q,h_2,\ \phi),$ $h_2 + \Pi_\varepsilon)dS(\phi \mid q,b)$. Since V_T is independent of the precision h in the last period, we need only consider the effect of changes in h on the q' function. Note that q' is a random variable, with

$$E(q' \mid q,h) = \frac{hq + \Pi_\varepsilon q}{h + \Pi_\varepsilon} = q$$

where expectation is taken with respect to $dS(\phi \mid q,b)$. The conditional variance of q' is given by

$$E\left((q' - q)^2 \mid q,h\right) = \frac{1}{h + \Pi_\varepsilon}.$$

Then decreases in h correspond to increases in the variance of the random variable q'. Since the mean of q' is not a function of h and q' is normally distributed, this implies decreases in h correspond to conducting a mean preserving spread on q'. The function V_T is convex increasing so we have shown $\int V_T(q'(q, h_1, \phi), h_1 + \Pi_\varepsilon)dS(\phi \mid q,b) < \int V_T(q'(q, h_2, \phi), h_2 + \Pi_\varepsilon)\, dS(\phi \mid q,b)$.

Now replace the variance b with the true variance of the random variable ϕ, $(h + \Pi_\varepsilon)/h\Pi_\varepsilon$. The function q' is increasing in ϕ, and of course V_T is convex increasing in q'. The partial derivative of the variance of ϕ with respect to h is $-1/h^2 < 0$. Then decreases in h amount to conducting a MPS directly on ϕ, which must increase the value of the problem. We have shown that decreases in h will increase the value of expression (2.6), and therefore V_{T-1} is decreasing in h. This analysis may be repeated inductively for V_{T-2}, V_{T-3}, ..., V_1, and we establish

$$V_t(q,h) = \max\{q + \beta\int V_{t+1}(q'(q,\ h,\ \phi),\ h')dS(\phi \mid q,\ \frac{h + \Pi_\varepsilon}{h\Pi_\varepsilon});$$

$$-C + \beta\int V_{t+1}(q'(0,\ \Pi_\theta,\ \phi),\ \Pi_\theta + \Pi_\varepsilon)\ dS(\phi \mid 0,\ \frac{\Pi_\theta + \Pi_\varepsilon}{\Pi_\theta \Pi_\varepsilon})\}$$

is increasing in q for t = 1, ..., T and is decreasing in h for t = 1, ..., T-1.

Using these monotonicity properties of V_t we can construct reservation value functions $q_t^*(h)$ for t = 1, ..., T. We have already discussed $q_T^*(h)$ above. These $q_t^*(h)$ are solutions of the implicit function

$$q_t^*(h) + \beta\int V_{t+1}(q'(q_t^*(h),\ h,\ \phi),\ h')dS(\phi \mid q_t^*(h),\ \frac{h + \Pi_\varepsilon}{h\Pi_\varepsilon})$$

$$= -C + \beta\int V_{t+1}(q'(0,\ \Pi_\theta,\ \phi),\ \Pi_\theta + \Pi_\varepsilon)dS(\phi \mid 0,\ \frac{\Pi_\theta + \Pi_\varepsilon}{\Pi_\theta \Pi_\varepsilon}).$$

Since V_t is decreasing in h, the workers at matches with large h must have higher values of q to induce them to stay at the firm. Then $q_t^*(h)$ is an increasing function. For a worker who has been with the current firm for d periods, $h = \Pi_\theta + d\Pi_\varepsilon$, so h is an increasing function of d. The longer the duration of the match, the higher must be the conditional mean q for the worker not to leave the match.

We have established the existence of reservation value rules of the form

(S) stay with current firm iff $q \geq q_t^*(h)$

(M) leave current firm iff $q < q_t^*(h)$, $t = 1, \ldots, T$.

By the boundedness assumptions and the existence of a discount factor $\beta \varepsilon (0,1)$, we have sufficient conditions for the optimal value function of the T horizon problem to converge to the optimal value of the infinite horizon problem. Write this function as

$$V(q,h) = \max\left\{ q + \beta \int V(q'(q,\ h,\ \phi),\ h') dS(\phi \mid q,\ \frac{h + \Pi_\varepsilon}{h\Pi_\varepsilon});\right.$$

$$\left. -C + \beta \int V(q'(0,\ \Pi_\theta,\ \phi),\ \Pi_\theta + \Pi_\varepsilon) dS(\phi \mid 0,\ \frac{\Pi_\theta + \Pi_\varepsilon}{\Pi_\theta \Pi_\varepsilon})\right\}.$$

As for the finite horizon problem, V is increasing in q and decreasing in h. There exists an associated stationary reservation value policy of the form

(S) stay with current firm iff $q \geq q^*(h)$

(2.7)

(M) leave current firm iff $q < q^*(h)$.

3. SAMPLE COMPOSITION AND DESCRIPTIVE STATISTICS

The data used in the empirical work which follows are from the National Longitudinal Survey of Young Men. Interviews of approximately 5000 young men (14-24 years of age in 1966) were made at yearly intervals beginning in 1966 and ending in 1971.[6] There are approximately 5000 sample members, who were randomly selected from the national population of this age and sex group.

Virtually all models of turnover, both theoretical and empirical, posit a relation between the current period mobility decision and past labor market experiences. In the matching model considered in this paper, the turnover decision rule contains as arguments statistics which are functions of the wage history on the current job. If only a portion of that history is actually observed, estimation of structural parameters is greatly complicated by the existence of initial conditions problems.[7] Differences in initial conditions across individuals may often be misinterpreted as reflecting individual differences in tastes and technology (which is conventionally defined as unobserved heterogeneity). The distinction between these two types of heterogeneity is important if we are to obtain consistent estimates of structural parameters.

To circumvent the initial conditions problem,[8] the subsample (hereafter referred to as the "sample") was selected in the following manner. As of the first wave of the panel (1966) all sample members were enrolled in full-time schooling and reported themselves as not being full-time members of the labor force. At each of the interview dates in 1967, 1968, and 1969, the sample members were holding full-time jobs and

Table 1

Characteristics of the Sample Members

Variable[1]	Mean	Standard Deviation
LW_{67}	.714	.451
LW_{68}	.898	.432
LW_{69}	.972	.390
$M_{67,68}$	.391	
$M_{68,69}$	.315	
S	12.391	2.533
Black	.25	
A_{67}	19.85	2.750

[1]Variables are described in the text.

were not enrolled in full-time schooling. Individuals who were unemployed at the time of the 1967-1969 interviews were excluded because we did not wish to simultaneously consider turnover and unemployment decisions. As of each interview date we know the individual's hourly wage rate at his current job, and we know whether the individual is employed at the same firm in 1968 as he was in 1967 and if he is employed at the same firm in 1969 as he was in 1968. This sequence of wage rates and mobility indicators constitute the endogenous variables of the analysis. Potential sample members were also excluded if they had missing observations on any of the endogenous or exogenous variables included in the study. The selection criteria employed are obviously stringent. Their usage presumably minimizes the effects of differences in initial conditions on the structural estimates obtained. On the other hand, we should be cautious in drawing inferences concerning the nature of the mobility and wage growth process for young workers from structural estimates using this sample.

There were 248 individuals who satisfied all of the selection criteria. Table 1 contains the means and standard deviations of the principal variables of interest in this analysis. The variables included in Table 1 are defined as follows: The logarithm of deflated (1967 dollars) hourly wage rates in 1967, 1968, and 1969 are denoted LW_{67}, LW_{68}, and LW_{69}, respectively. If the individual changed firms between the 1967 and 1968 interview dates, $M_{67,68} = 1$, and $M_{67,68} = 0$ if he did not change employers. The indicator variable $M_{68,69}$ is similarly defined for the period between the 1968 and 1969 interviews. The number of years of schooling completed as of 1967 is denoted by S. No sample members are engaged in full-time schooling during the sample period so S is constant for each individual. The indicator Black is equal to 1 if the respondent is black and is equal to 0 if he is white. Finally, A_{67} is the age of the individual in 1967.

Table 2 contains the means and standard deviations of deflated hourly wages in levels for the entire sample and for each of the four mobility groups. It is apparent that wage levels and wage change greatly differ across mobility groups. In all mobility groups the increase in average wages between the second and third year is small relative to the change in average wages between the first and second years. Individuals who change employers between each pair of successive interviews have by far the lowest average wage of any of the groups and their change in average wage between 1967 and 1968 is enormous. Workers who move between 1967 and 1968 and then stay at the same employer in 1968 and 1969 have large wage gains when they move and a very small average change when they stay. The group of individuals that stay with the same employer between 1967 and 1968 and then move between 1968 and 1969 experience small wage gains between 1967-68 and 1968-69. Workers who remain with the same employer all three periods also have small average wage change increases, though their average first period wage rate is over one dollar greater than for workers who change employers twice. It appears that interfirm mobility tends to reduce the relative wage inequality of recent labor market entrants.

Table 2

Means and Standard Deviations of Hourly Wages by Mobility Status

Mobility Status	Mean	Standard Deviation	N
Entire Sample			
W_{67}	2.24	.905	248
W_{68}	2.67	1.093	248
W_{69}	2.84	1.051	248
Move 67; Move 68			
W_{67}	1.55	.585	42
W_{68}	2.20	1.236	42
W_{69}	2.27	.798	42
Move 67; Stay 68			
W_{67}	2.01	.785	55
W_{68}	2.69	1.018	55
W_{69}	2.88	.946	55
Stay 67; Move 68			
W_{67}	2.28	.965	36
W_{68}	2.48	1.032	36
W_{69}	2.72	1.013	36
Stay 67; Stay 68			
W_{67}	2.58	.866	115
W_{68}	2.89	1.038	115
W_{69}	3.06	1.119	115

The model described in Section 2 implies a number of restrictions on the form of the dynamic wage equations and the turnover decision rule. The following sections discuss estimation of structural parameters using our theoretical formulation, but at this point it seems desirable to test some of these restrictions using a fairly general econometric framework. Consider the following two-equation system:

$$\ell n\ w_{i1} = \alpha_1^M(m_i(1-p_i)) + \alpha_1^S((1-m_i)(1-p_i)$$

$$+ (\underset{\sim}{Z}_{i1}(m_i(1-p_i)))'\underset{\sim}{\gamma}_1^M + (\underset{\sim}{Z}_{i1}((1-m_i)(1-p_i)))'\underset{\sim}{\gamma}_1^S + \xi_{i1} \tag{3.1}$$

$$\ell n\ w_{i2} = \alpha_2^M(m_ip_i) + \alpha_2^S((1-m_i)p_i)$$

$$+ (\underset{\sim}{Z}_{i2}(m_ip_i))'\underset{\sim}{\gamma}_2^M + (\underset{\sim}{Z}_{i2}((1-m_i)p_i))'\underset{\sim}{\gamma}_2^S + \xi_{i2}$$

where m_i = 1 iff i changed firms between t = 1 and t = 2

= 0 else

p_i = 1 iff t = 2 (year 1968)

= 0 else (year 1967)

α_k^j is the intercept term for mobility group j in year k,

$\underset{\sim}{\gamma}_k^j$ is the coefficient vector for mobility group j in year k,

and (ξ_{i1}, ξ_{i2}) are disturbances. While the random variables η, θ, and ε were assumed to be normally distributed in Section 2, ξ_{i1}, and ξ_{i2} will not be normally distributed after selection on θ and ε has taken place. Thus we test the restrictions of the model using the minimum distance estimator described in Chamberlain (1982). We can allow for general forms of heteroskedasticity in (ξ_{i1}, ξ_{i2}) using this estimator, and no explicit distribution assumptions on (ξ_{i1}, ξ_{i2}) are necessary to construct asymptotically valid tests of the restrictions. The tests were conducted under two sets of assumptions on 3.1,

(3.2)

A. i. $E(\underset{\sim}{\ln w_i} \mid \underset{\sim}{X}_i) = \underset{\sim}{X}_i' \underset{\sim}{\delta}$
(linear conditional expectation)

ii. $\mathrm{Var}(\underset{\sim}{\ln w_i} \mid \underset{\sim}{X}_i) = \Sigma$
(constant covariance matrix)

B. i. $P(\underset{\sim}{\ln w_i} \mid \underset{\sim}{X}_i) = \underset{\sim}{X}_i' \delta \neq E(\underset{\sim}{\ln w_i} \mid \underset{\sim}{X}_i)$
(linear predictor, nonlinear conditional expectation)

ii. $$\Omega = E\{(\underset{\sim}{\ln w_i} - \underset{\sim}{X}_i'\underset{\sim}{\delta})(\underset{\sim}{\ln w_i} - \underset{\sim}{X}_i\underset{\sim}{\delta})' [E(\underset{\sim}{X}_i\underset{\sim}{X}_i')]^{-1}(\underset{\sim}{X}_i\underset{\sim}{X}_i')]E(\underset{\sim}{X}_i\underset{\sim}{X}_i')]^{-1}\}$$

where $\Omega \equiv$ asymptotic covariance matrix of $\sqrt{I}\,(\hat{\underset{\sim}{\delta}} - \underset{\sim}{\delta})$, I is the number of individuals, $\underset{\sim}{\ln w_i} = (\ln w_{i1}\ \ln w_{i2})'$,

$$\underset{\sim}{X}_i' = \begin{bmatrix} \underset{\sim}{X}_{i1}' \\ \underset{\sim}{X}_{i2}' \end{bmatrix}$$

$\underset{\sim}{X}_{it}$ is a $k \times 1$ vector of observed heterogeneity components at time t, and $\underset{\sim}{\delta}$ is a conformable parameter matrix.

We tested several linear restrictions on coefficients in $\underset{\sim}{\delta}$ using (3.1) under both sets of assumptions A and B in (3.2). Since both sets of assumptions resulted in the same inference concerning rejection or nonrejection of restrictions, we report results of tests using the assumptions contained in A. The general strategy used in nesting the tests was to first test restrictions on the coefficient vector $\underset{\sim}{\gamma}_k^j$ and then move to the intercept terms α_k^j.

We first tested the restriction that the coefficient vectors for movers and stayers were time invariant, though they were allowed to depend on mobility status. This restriction appears in line 1 of Table 3. The minimum distance statistic produced does not lead to strong rejection of the restriction. Having not rejected this restriction, we proceed to test a more basic implication of the matching model. Since $\underset{\sim}{Z}_{it}$ does not enter the turnover decision rule, there is no reason to expect $\underset{\sim}{\gamma}^M \neq \underset{\sim}{\gamma}^S$. This restriction also produced a relatively small distance statistic. There is no strong empirical evidence to suggest that the coefficient vector $\underset{\sim}{\gamma}$ is not time invariant, and conditional on time invariance, is not independent of mobility status.

We then proceed to test restrictions on the intercept terms. By the theory, the intercept terms for the firm-changers should increase substantially between periods one and two. The intercept term for individuals who do not change may also increase through the accumulation of general human capital. In line 4 of Table 3 we impose the restriction of time invariance of the intercept terms of the mover and stayer groups. This restriction is overwhelmingly rejected. In lines 3a and 3b we test the time invariance restrictions one at a time. Although the restriction is violated for both movers and stayers, rejection is by far the clearest for the group of movers. Lastly, in line 5, we test equality of α^M and

Table 3

Tests of Nested Restrictions in Linear Model

Restrictions	χ^2	df	$\chi^2_{.95}(df)$
1. $\underset{\sim}{\gamma}_1^M = \underset{\sim}{\gamma}_2^M = \underset{\sim}{\gamma}^M$ $\underset{\sim}{\gamma}_1^S = \underset{\sim}{\gamma}_2^S = \underset{\sim}{\gamma}^S$	10.98	6	12.6
2. $\underset{\sim}{\gamma}^M = \underset{\sim}{\gamma}^S = \underset{\sim}{\gamma}$	4.57	3	7.81
3a. $\alpha_1^M = \alpha_2^M = \alpha^M$	67.09*	1	3.84
3b. $\alpha_1^S = \alpha_2^S = \alpha^S$	7.19*	1	3.84
4. $\begin{bmatrix} \alpha_1^M = \alpha_2^M = \alpha^M \\ \alpha_1^S = \alpha_2^S = \alpha^S \end{bmatrix}$	68.27*	2	5.99
5. $\alpha^M = \alpha^S$	58.94	1	3.84

*Using model (2) as baseline.

α^S conditional on time invariance of the intercepts. This restriction is also decisively rejected.

We conclude this section by considering the sensitivity of parameter estimates with respect to alternative assumptions regarding the linearity of the conditional expectation function and heteroskedasticity of the disturbances. Equation (3.1) was estimated under assumptions A and B in (3.2) and imposing time and mobility status invariance on $\underset{\sim}{\gamma}_k^j$. The results appear in Table 4. It seems reasonable to claim that the parameter estimates are relatively insensitive to the assumptions made. For the purpose of obtaining point estimates, nonlinearity of the conditional expectation function and heteroskedasticity do not appear to be important considerations in this particular application.

4. INITIAL CONSISTENT ESTIMATES OF STRUCTURAL PARAMETERS

All structural parameters which are potentially identifiable can be consistently estimated using only the first <u>two</u> years of data for the sample described in the previous section. To reduce the computational burden (particularly in obtaining maximum likelihood estimates), we will restrict our analysis to these two periods. Obtaining initial consistent estimates can serve at least two useful purposes for this analysis. First, they constitute good starting values for the subsequent maximum likelihood estimation. It is also the case that taking one Newton step from the initial consistent estimates produces estimates which are asymptotically efficient. Thus for computationally demanding models we may wish to perform only one iteration to achieve asymptotically efficient

Table 4

Estimates of Log Wage Equations under
Assumptions (A) and (B) of Equation 3.2

Parameters	A	B
α_1^M	-.506 (2.892)*	-.530 (3.180)
α_2^M	-.286 (1.539)	-.330 (1.896)
α_1^S	-.204 (1.165)	-.260 (1.550)
α_2^S	-.178 (.956)	-.221 (1.279)
School	.033 (2.520)	.030 (2.421)
Black	-.134 (2.419)	-.153 (3.147)
Age_t	.036 (2.912)	.040 (3.394)

*Absolute value of normal statistic in parentheses.

estimates. Second, it is of interest to determine what restrictions the matching model imposes on moments of least squares residuals from log wage equations for various mobility groups and time periods. In particular, some interesting issues concerning the identification of σ_θ^2 and σ_ε^2 arise.

It is straightforward to obtain initial consistent estimates of the parameter vector $\underset{\sim}{\gamma}$. In the specification considered here, $\underset{\sim}{\gamma}$ includes an intercept term, the number of years of schooling completed, the indicator variable for the race of the individual, and the individual's age. We can obtain consistent estimates of $\underset{\sim}{\gamma}$ from the regression of the log of the first period wage rate on these regressors. No systematic selection has as yet taken place, and the estimate of the intercept term is unbiased. The results of the OLS regression are:

$$\underset{}{LW_{67}} = \underset{(3.629)}{-.683} + \underset{(2.176)}{.031\ S} - \underset{(2.239)}{.134\ \text{Black}} + \underset{(4.047)}{.052\ A_{67}}$$

$$R^2 = .262,$$

where the absolute value of the t-statistic appears in parentheses.

We make all the assumptions concerning the first and second moments and the orthogonality of η, θ, and ε that were made in Section 2. We immediately proceed to a specification of the first and second period wage equations for movers and stayers. The log wage equations for job changers are

$$w_{ij1} = \underset{\sim}{Z}_{i1}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij} + \varepsilon_{ij1}; \ \theta_{ij} + \varepsilon_{ij1} \leq q_1^*$$

$$w_{ij'2} = \underset{\sim}{Z}_{i2}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij'} + \varepsilon_{ij'2}.$$

The log wage equations for stayers are

$$w_{ij1} = \underset{\sim}{Z}_{i1}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij} + \varepsilon_{ij1}$$

$$w_{ij2} = \underset{\sim}{Z}_{i2}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij} + \varepsilon_{ij2}, \ \theta_{ij} + \varepsilon_{ij1} \geq q_1^*.$$

Again, by the orthogonality assumptions, ordinary least squares (OLS) estimates from the first period log wage equation are consistent estimates of $\underset{\sim}{\gamma}$. An estimate of total variation $\sigma_T^2 = \sigma_\eta^2 + \sigma_\theta^2 + \sigma_\varepsilon^2$ is given by

$$(4.1) \quad \hat{\sigma}_T^2 = \frac{1}{I} \sum_{i=1}^{I} (w_{ij1} - \underset{\sim}{Z}_{i1}'\hat{\underset{\sim}{\gamma}})^2,$$

where $\hat{\underset{\sim}{\gamma}}$ is the vector of consistent OLS estimates of $\underset{\sim}{\gamma}$. For this sample, $\hat{\sigma}_T^2 = .150$.

We can obtain an estimator of the sum $\sigma_\theta^2 + \sigma_\varepsilon^2$ in the following manner. For all movers ($i \varepsilon M$), change in log wage is

$$(w_{ij'2} - w_{ij1} \mid i \varepsilon M) = (\underset{\sim}{Z}_{i2} - \underset{\sim}{Z}_{i1})'\underset{\sim}{\gamma} + \theta_{ij'} - \theta_{ij} + \varepsilon_{ij'2} - \varepsilon_{ij1};$$

$$\theta_{ij} + \varepsilon_{ij1} \leq q_1^*.$$

Then consider

$$(4.2)\quad E[(w_{ij'2} - w_{ij1} - (\underset{\sim}{Z}_{i1} - \underset{\sim}{Z}_{i1})'\underset{\sim}{\gamma})(w_{ij'2} - \underset{\sim}{Z}_{i2}'\underset{\sim}{\gamma}) \mid i\varepsilon M] =$$

$$E[(\theta_{ij'} - \theta_{ij} + \varepsilon_{ij'2} - \varepsilon_{ij1})(\eta_i + \theta_{ij'} + \varepsilon_{ij'2}) \mid \theta_{ij} + \varepsilon_{ij1} < q_1^*]$$

$$= E[\theta_{ij'}^2 + \varepsilon_{ij'2}^2 \mid \theta_{ij} + \varepsilon_{ij1} < q_1^*]$$

$$= E[\theta_{ij'}^2 + \varepsilon_{ij'2}^2]$$

$$= \sigma_\theta^2 + \sigma_\varepsilon^2.$$

After replacing γ with $\hat{\gamma}$ in (4.2), we get that an initial consistent estimate of $\sigma_\theta^2 + \sigma_2^2$ is

$$(4.3)\quad (\sigma_\theta^2 + \sigma_\varepsilon^2) = \frac{1}{\#(M)} \sum_{i\varepsilon M} (w_{ij'2} - w_{ij1} - (\underset{\sim}{Z}_{i1} - \underset{\sim}{Z}_{i1})'\hat{\underset{\sim}{\gamma}})(w_{ij'2} - \underset{\sim}{Z}_{i2}'\hat{\underset{\sim}{\gamma}})$$

Using these two estimates, an estimate of σ_η^2 is

$$(4.4)\quad \hat{\sigma}_\eta^2 = \hat{\sigma}_T^2 - (\sigma_\theta^2 + \sigma_\varepsilon^2)$$

We obtain estimates of $\sigma_\theta^2 + \sigma_\varepsilon^2$ and σ_η^2 of .092 and .058, respectively. We now turn to the issue of identifying σ_θ^2 and σ_ε^2 separately.

Due to the information structure of the model, it is not possible to separately identify σ_θ^2 and σ_ε^2 without the imposition of specific distributional assumptions. In the language of Flinn and Heckman (1982b), σ_θ^2 and σ_ε^2 are nonparametrically underidentified. Since the matching model is developed for the case in which the stochastic terms θ and ε are normally distributed, we adopt this assumption at this point.

The second period log wage equation for stayers has expected value

$$E(w_{ij2} \mid \underset{\sim}{Z}_{i2}', i\varepsilon S) = \underset{\sim}{Z}_{i2}'\underset{\sim}{\gamma} + E(\theta_{ij} + \varepsilon_{ij2} \mid \theta_{ij} + \varepsilon_{ij1} \geq q_1^*).$$

But, since ε is i.i.d.,

$$(4.5) \quad E(\theta_{ij} + \varepsilon_{ij2} \mid \theta_{ij} + \varepsilon_{ij1} \geq q_1^*) = E(\theta_{ij} \mid \theta_{ij} + \varepsilon_{ij1} \geq q_1^*)$$

$$= E_\varepsilon E(\theta_{ij} \mid \theta_{ij} \geq q_1^* - \varepsilon)$$

$$= \int_{-\infty}^{\infty} \frac{\phi(u)}{1-\Phi(u)} \frac{\sigma_\theta}{\sigma_\varepsilon} \phi(\frac{\varepsilon}{\sigma_\varepsilon})d\varepsilon$$

where $u = (q_1^* - \varepsilon)/\sigma_\theta$, $\phi(\cdot)$ is the p.d.f. of a standard normal random variable, and $\Phi(\cdot)$ is the c.d.f. of a standard normal r.v.

The residual from the second period log wage equation for stayers provides a consistent estimator of the LHS of (4.5). The expression on the RHS of (4.5) involves σ_ε, σ_θ, and q_1^*. One of the standard deviations is redundant, however, since we have an estimate of $\sigma_\theta^2 + \sigma_\varepsilon^2$, which we denote by σ_v^2. Then $\sigma_\varepsilon = (\sigma_v^2 - \sigma_\theta^2)^{1/2}$. From our OLS regression residuals it is possible to construct two consistent estimates of q_1^*. For movers in period 1, the expectation of the disturbance term is

$$(4.6)\quad E(\theta_{ij} + \varepsilon_{ij1} \mid \theta_{ij} + \varepsilon_{ij1} \leqslant q_1^*) = -\frac{\phi(\frac{q_1^*}{\sigma_v})}{\Phi(\frac{q_1^*}{\sigma_v})}\sigma_v.$$

For stayers, the expected value of the disturbance term in the first period log wage equation is

$$(4.7)\quad E(\theta_{ij} + \varepsilon_{ij1} \mid \theta_{ij} + \varepsilon_{ij1} \geqslant q_1^*) = \frac{\phi(\frac{q_1^*}{\sigma_v})}{1-\Phi(\frac{q_1^*}{\sigma_v})}\sigma_v.$$

The LHS of (4.6) and (4.7) are consistently estimated by the mean of first period OLS residuals for the mover and stayer groups, respectively. Since a consistent estimate of σ_v is available, it is possible to solve (4.6) and (4.7) for two estimates of q_1^* (both consistent under the normality assumption). The estimates of q_1^* are .261 and -.294 for movers and stayers, respectively.

Using consistent estimates of q_1^* and σ_v^2, we may solve (4.5) for estimates of σ_θ^2 and σ_ε^2. In general this procedure does not yield unique estimates of σ_θ^2 and σ_ε^2, even for a given value of q_1^*. With no clear criteria by which we may choose between these estimates, this procedure is mainly useful for computational purposes. Any one of these estimates may be used as starting values to guarantee consistency of the maximum likelihood estimator described in the next section.

5. MAXIMUM LIKELIHOOD ESTIMATES OF THE MATCHING MODEL

In this section we present maximum likelihood estimates of the identifiable structural parameters. In general, we are not able to identify the mobility cost parameter C and the discount factor β. While it is possible to map out a locus of equiprobable pairs (C, β), there exists no extraneous information which allows us to choose a particular pair (although we would want to restrict estimates to the region of the plane in which $C \geq 0$, $0 \leq \beta < 1$). It is possible to compute fully efficient maximum likelihood estimates for $(\underset{\sim}{\gamma}', \sigma_\eta^2, \sigma_\theta^2, \sigma_\varepsilon^2) \equiv \underset{\sim}{\Gamma}$ in a relatively straightforward manner. We will only concern ourselves with estimating the parameter vector $\underset{\sim}{\Gamma}$ in this paper.

The stochastic structure is assumed to be of the form:

$$\eta_i \sim N(0, \sigma_\eta^2) \qquad \forall i$$

$$\theta_{ij} \sim N(0, \sigma_\theta^2) \qquad \forall i,j$$

$$\varepsilon_{ijt} \sim N(0, \sigma_\varepsilon^2) \qquad \forall i,j,t,$$

and all stochastic components are assumed to be pairwise orthogonal for all subsets of relevant indices from $\{i,j,t\}$.

Once again we use only the first two periods of observations. The dependent variables of the analysis consist of log wages in periods one and two and the interfirm mobility indicator for the 1967-68 period. The joint probability function of wages and mobility conditional on observed heterogeneity is $k(w_{ij1}, w_{ij'2}, m_{i2} \mid \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \underset{\sim}{\Gamma})$. Rewrite the joint probability function as

$$(5.1) \quad k(w_{ij1}, w_{ij'2}, m_{i2} \mid \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \underset{\sim}{\Gamma}) = k_1(w_{ij1}, w_{ij'2} \mid m_{i2}, \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \underset{\sim}{\Gamma}) \times k_2(m_{i2} \mid \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \underset{\sim}{\Gamma})$$

where k_1 is the bivariate density of log wages conditional on mobility status and k_2 is the marginal probability of mobility.

By the reservation value property established in Section 2 and the information structure of the problem, we know that observed turnover between the first and second period depends on whether $\theta_{ij} + \varepsilon_{ij1} \gtrless q_1^*$, or $\varepsilon_{ij1} \gtrless Q^* \equiv q_1^* - \theta_{ij}$. Then, conditional on unobserved heterogenity components θ_{ij} and η_i, we can write

$$(5.2) \quad \ell_1(w_{ij1}, w_{ij'2} \mid m_{i2}, \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \eta_i, \theta_{ij}, \Gamma) = \ell_1^{(1)}(w_{ij1} \mid m_{i2}, \underset{\sim}{Z}_{i1}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma}) \times \ell_1^{(2)}(w_{ij'2} \mid m_{i2}, \underset{\sim}{Z}_{i2}, \eta_i, \theta_{ij1}, \underset{\sim}{\Gamma})$$

and

$$(5.3) \quad \ell_2(m_{i2} \mid \underset{\sim}{Z}_{i1}, \underset{\sim}{Z}_{i2}, \eta_i, \theta_{ij}, \Gamma) = \ell_2(m_{i2} \mid \theta_{ij}, \underset{\sim}{\Gamma}).$$

The integrated likelihood function (ILF) may be written as the product of the ILFs for the sets of movers and stayers. The sample ILF is given by

$$(5.4) \qquad \bar{L}(\underset{\sim}{\Gamma}) = \bar{L}_M(\underset{\sim}{\Gamma})\bar{L}_S(\underset{\sim}{\Gamma})$$

where

$$(5.5) \qquad \bar{L}_M(\underset{\sim}{\Gamma}) = \prod_{i\varepsilon M} \int_{\Omega_\eta} \Big[\int_{\Omega_\theta} \frac{\ell_1^{(1)}(w_{ij1} \mid m_{i2}=1, \underset{\sim}{Z}_{i1}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma})}{F(q_1^* - \theta)} F(q_1^* - \theta)dG(\theta_{ij})\Big]$$

$$\times\ \ell_1^{(2)}(w_{ij'2} \mid m_{i2}=1, \underset{\sim}{Z}_{i2}, \eta_i, \underset{\sim}{\Gamma})dH(\eta_i)$$

and

$$(5.6) \qquad \bar{L}_S(\underset{\sim}{\Gamma}) = \prod_{i\varepsilon S} \int_{\Omega_\eta} \Big[\int_{\Omega_\theta} \frac{\ell_1^{(1)}(w_{ij1} \mid m_{i2}=0, \underset{\sim}{Z}_{i1}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma})}{1-F(q_1^* - \theta)}$$

$$\times\ \ell_1^{(2)}(w_{ij2} \mid m_{i2}=0, \underset{\sim}{Z}_{i2}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma})(1 - F(q_1^* - \theta_{ij}))\Big]\, dG\,(\theta_{ij})dH(\eta_i).$$

In the ILF for the movers (5.5), the conditional density of first period wages has a truncated normal form. The conditional density $\ell_1^{(1)}(w_{ij}, \mid m_{i2}=1, \underset{\sim}{Z}_{i1}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma})/F(q_1^* - \theta)$ is truncated normal with untruncated mean and variance equal to $\underset{\sim}{Z}_{i1}'\underset{\sim}{\gamma} + \eta_i + \theta_{ij}$ and σ_ε^2, respectively, and upper truncation point $q_1^* - \theta$. When we multiply by the marginal probability of turnover which is simply $F(q_1^* - \theta)$, we get

$$(5.5)' \quad \bar{L}_M(\underset{\sim}{\Gamma}) = \prod_{i\varepsilon M} \int_{\Omega_\eta} \left[\int_{\Omega_\theta} \ell_1^{(1)}(w_{ij1} \mid m_{i2}=1, \underset{\sim}{Z}_{i1}, \eta_i, \theta_{ij}, \underset{\sim}{\Gamma}) dG(\theta_{ij})\right]$$

$$\times \ \ell_1^{(2)}(w_{ij'2} \mid m_{i2} = 1, \underset{\sim}{Z}_{i2}, \eta_i, \underset{\sim}{\Gamma}) dH(\eta_i).$$

Now integration of $\ell_1^{(1)}$ with respect to θ is unnecessary and we can rewrite $\ell_1^{(1)}$ as

$$\ell_1^{(1)}(w_{ij1} \mid m_{i2}=1, \underset{\sim}{Z}_{i1}, \eta_i, \underset{\sim}{\Gamma}),$$

which is a normal density with mean $\underset{\sim}{Z}_{i1}'\underset{\sim}{\gamma} + \eta_i$ and variance $\sigma_\theta^2 + \sigma_\varepsilon^2$. So that the final form of (5.5)' is

$$(5.5)'' \quad \bar{L}_M(\Gamma) = \int_{\Omega_\eta} \ell_1^{(1)}(w_{ij1} \mid m_{i2}=1, \underset{\sim}{Z}_{i1}, \eta_i, \underset{\sim}{\Gamma})$$

$$\times \ \ell_i^{(2)}(w_{ij'2} \mid m_{i2}=1, \underset{\sim}{Z}_{i2}, \eta_i, \underset{\sim}{\Gamma}) dH(\eta_i).$$

In a similar manner we can simplify the ILF contribution of stayers (5.6). Once again the conditional density function of first period log wages is truncated normal, with <u>lower</u> truncation point $q_1^* - \theta$. After multiplying by the probability of staying $\left(1 - F(q_1^* - \theta)\right)$, we get

$$(5.6)' \quad \bar{L}_S(\underset{\sim}{\Gamma}) = \prod_{i\varepsilon S} \int_{\Omega_\eta} \left[\int_{\Omega_\theta} \ell_1^{(1)}(w_{ij1} \mid m_{i2}=0,\ \underset{\sim}{Z}_{i1},\ \eta_j,\ \theta_{ij},\ \underset{\sim}{\Gamma})\right.$$

$$\left. \times\ \ell_1^{(2)}(w_{ij2} \mid m_{i2} = 0,\ \underset{\sim}{Z}_{i2},\ \eta_i,\ \theta_{ij},\ \underset{\sim}{\Gamma})dG(\theta_{ij})\right] dH(\eta_i)$$

Now define the random variable $\delta_{ij} \equiv \eta_i + \theta_{ij}$, so that $\delta_{ij} \sim N(0,\ \sigma_\eta^2 + \sigma_\theta^2)$. Then we may rewrite $(5.6)'$ as

$$(5.6)'' \quad \bar{L}_S(\underset{\sim}{\Gamma}) = \prod_{i\varepsilon S} \int_{\Omega_\delta} \ell_1^{(1)}(w_{ij'} \mid m_{i2} = 0,\ \underset{\sim}{Z}_{i1},\ \delta_{ij},\ \underset{\sim}{\Gamma})$$

$$\times\ \ell_1^{(2)}(w_{ij2} \mid m_{i2} = 0,\ \underset{\sim}{Z}_{i2},\ \delta_{ij},\ \underset{\sim}{\Gamma})dD(\delta_{ij}),$$

where D is the cumulative distribution function of δ_{ij}. Both $(5.5)''$ and $(5.6)''$ require performing only one numerical integration per observation per iteration in performing the nonlinear optimization required for the estimation of $\underset{\sim}{\Gamma}$.

Let $\bar{\mathcal{L}}(\underset{\sim}{\Gamma}) = \ln\bar{L}(\underset{\sim}{\Gamma}) = \ln\bar{L}_M(\underset{\sim}{\Gamma}) + \ln\bar{L}_S(\underset{\sim}{\Gamma})$. Maximum likelihood estimates of $\underset{\sim}{\Gamma}$ were obtained by maximizing $\bar{\mathcal{L}}(\underset{\sim}{\Gamma})$ with respect to $\underset{\sim}{\Gamma}$ using a modified method of scoring technique. The maximum likelihood estimates and asymptotic standard normal statistics are presented in Table 5. In comparison with the consistent estimates of $\underset{\sim}{\gamma}$ which were used as starting values in the maximization of $\bar{\mathcal{L}}(\Gamma)$, only the OLS and ML estimates of the schooling coefficient differ appreciably. The ML estimate of the schooling coefficient is extremely small and only marginally significant. The ML

Table 5

Maximum Likelihood Estimates of Jovanovic-Type Matching Model

Parameter	Point Estimate	Asymptotic Normal Statistic
Constant	-.766	-8.014
S	.011	1.835
Black	-.145	-4.355
Age_t	.073	14.028
σ_η^2	.079	4.933
σ_θ^2	.068	5.686
σ_ε^2	.031	13.228
£	-150.082	

estimate of the age coefficient is large relative to the OLS estimate and highly significant.

Given the structure of the stochastic term assumed in this paper, it is possible to assess the proportion of variability in log wages of new labor market entrants due to η, θ, and ε. In the first period, .44 of the total variability is attributable to individual specific, time and job invariant heterogeneity, .38 is due to worker-firm specific heterogeneity, and the remainder to white noise. As a cohort (defined in terms of date of labor market entrance) ages, the proportion of log wage X variability attributable to worker-firm heterogeneity will decline as individuals sort into acceptable matches.[9] Thus .38 represents an upper bound on the proportion of variance attributable to match specific factors.

6. CONCLUSION

We have demonstrated that worker-firm matching heterogeneity is an important factor in the explanation of wage variation among recent labor market entrants. This suggests that the stochastic structure of earnings may differ substantially between groups characterized by different mobility patterns. Econometric models of earnings dynamics which do not incorporate the mobility decision may be seriously misspecified if the sample includes a significant proportion of recent labor market entrants.

We should not overstate the importance of matching heterogeneity as an explanation of turnover. When a firm is first sampled, the individual's prior on θ_{ij} has variance .068. After one period (one productivity realization), the variance associated with the estimate of θ is reduced to .021. After two realizations, it is further reduced to .013. Most of the uncertainty as to the value of the match is resolved in the first two periods of employment. Thus, matching heterogeneity may be an important explanation of terminations of brief worker-firm attachments, but does not promise to be useful in explaining separations at tenure levels greater than three or four years, at least as currently formulated.

Footnotes

[1]This assumption is not strictly necessary, but is made for simplicity. Johnson makes the assumption that agents only live for two periods, which is overly restrictive for our purposes.

[2]The demand side of the market cannot be addressed empirically, because virtually all currently available data sets are supply side oriented, i.e., the unit of analysis is the individual.

[3]The division of these rents is obviously an important issue, which has recently been addressed by Mortensen (1978).

[4]For the sample of young men I use in the empirical work reported in Section 2, between 1968 and 1969 only 15% of all workers changing jobs indicated that they had involuntarily left their last job.

[5]If the random variable y is normally distributed with mean 0 and variance σ_y^2, its support is the real line $(-\infty, \infty)$. Denote the p.d.f. and c.d.f. of y by f(y) and F(y) respectively. We can then create the transformed variable y^* with p.d.f. defined by

$$f^*(y^*) = \frac{f(y^*)}{F\left(\frac{M}{\sigma_y}\right) - F\left(\frac{-M}{\sigma_y}\right)} \qquad -M \leqslant y^* \leqslant M$$

$$= 0 \qquad y^* < -M;\ y^* > M$$

so that y^* is a truncated normal random variable with support $[-M, M]$, $0 < M < \infty$. Denote the c.d.f. of y^* by F^*. Define $D_M(x) = 1$ if $x \varepsilon [-M, M]$, $D_M(x) = 0$ if $x \notin [-M, M]$. Then $\lim_{M\to\infty} \sup_x \left| D_M(x)(F^*(x) - F(x)) \right| = 0$. We may choose M arbitrarily large so that the denominator in the expression for $f^*(y^*)$, $-M \leqslant y^* \leqslant M$, is approximately 1. For these reasons we may safely ignore any truncation correction in what follows.

[6]Interviews were also completed in subsequent years, but these occurred at intervals greater than one year and tended to be less comprehensive.

[7]For a discussion of the ramifications of initial conditions problems in the context of two different econometric models for the analysis of panel data, see Heckman (1981), and Flinn and Heckman (1982a).

[8]Of course the time at which an individual leaves full-time schooling and enters the labor market is determined by market conditions and personal characteristics. This endogeneity is ignored in our "solution" to the initial conditions problem.

[9]Given that θ and ε are normally distributed.

References

Becker, G.S. "Investment in Human Capital: A Theoretical Analysis." Journal of Political Economy 70, No. 5, Pt. 2 (October 1962): 9-49.

Burdett, K. "A Theory of Employee Job Search and Quit Rates." American Economic Review 68 (March 1978): 212-220.

Chamberlain, G. "Panel Data." S.S.R.I. Working Paper No. 8209, University of Wisconsin-Madison, April 1982. Forthcoming in the Handbook of Econometrics (Z. Griliches and M. Intrilligator, eds.).

DeGroot, M.H., Optimal Statistical Decisions. New York: McGraw-Hill, 1970.

Flinn, C. and J. Heckman. "Models for the Analysis of Labor Force Dynamics," in Advances in Econometrics, Vol. 1 (R. Basmann and G. Rhodes, eds.). Greenwich, Conn.: JAI Press, 1982a.

Flinn, C. and J. Heckman. "New Methods for Analyzing Structural Models of Labor Force Dynamics." Journal of Econometrics 18 (1982b): 115-168.

Heckman, J. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in Structural Analysis of Discrete Data (C. Manski and D. McFadden, eds.). Cambridge: MIT Press, 1981.

Johnson, W.R. "A Theory of Job Shopping." Quarterly Journal of Economics, 92, No. 2 (May 1978): 261-277.

Jovanovic, B. "Job Matching and the Theory of Turnover." Ph.D. Dissertation, University of Chicago, 1978.

Kuratani, M. "Schooling, Human Capital, and Labor Turnover." Ph.D. Dissertation, Columbia University, 1973.

Lippman, S. and J. McCall. "The Economics of Belated Information." Western Management Science Institute Working Paper No. 186, University of California at Los Angeles, December 1978.

Mortensen, D. "Specific Capital and Labor Turnover." Bell Journal of Economics 9, No. 2 (Autumn 1978): 572-586.

Parsons, D. "Specific Human Capital: An Application to Quit Rates and Layoff Rates." Journal of Political Economy 80, No. 6 (December 1972): 1120-1143.

Rosen, S. "Short-run Employment Variation on Class-I Railroads in the United States." Econometrica, 36, Nos. 3-4 (July-October 1968): 511-529.

Silcock, H. "The Phenomenon of Labor Turnover." Journal of the Royal Statistical Society, Series A 117, No. 3 (September 1954): 506-530.

Wilde, L. "An Information-theoretic Approach to Job Quits," in Studies in the Economics of Search (S. Lippman and J. McCall, eds.) Amsterdam: North Holland, 1980.

Wilson, C. "A Model of Job Search and Matching." Department of Economics, University of Wisconsin-Madison, February 1980.