Discussion Paper No. 630

QUIT PROBABILITIES AND JOB TENURE:

ON THE JOB TRAINING OR MATCHING?

Dale T. Mortensen

Northwestern University

August, 1984

Financial support from the U. S. Department of Labor is acknowledged. The results and interpretations reported in this draft are preliminary and, therefore, judgment should be used when quoting. Comments on the draft would be appreciated.

0. INTRODUCTION

A positive cross section association between the wage earned and length of service on a job, tenure, is a common observation reported in the empirical literature on mobility and earnings. One economic explanation for the phenomena is investment in job specific human capital. The theoretical literature on the subject identifies three forms of job specific investment that can account for the relationship -- job search, training on the job, and job matching. According to the first version of the hypothesis, an individual worker's wage is stationary on a specific job out differs across jobs. Information about the location and nature of jobs offering the different earnings is imperfect and takes time to acquire. Hence, a worker earning a higher wage relative to alternatives is less likely to quit. In the second and best known version of the hypothesis, a worker acquires job and/or employer specific skills through learning and training on the job. Consequently, the worker's wage increases with tenure relative to offers on alternative jobs which implies that the propensity to separate falls with tenure. In the third version, a worker's productivity on a specific job is initially uncertain but the degree of uncertainty diminishes with experience on the job as a consequence of observation. "Better" matches, those on which the worker experiences higher wage growth, endure.

Given the pure search model, a positive wage-tenure relation is a statistical artifact. It is the consequence of sample selection induced by the mobility process, movement by individual workers from lower to higher paying jobs at frequencies that diminish with the wage earned. Recent empirical evidence based on panel data attribute a significant proportion but not all wage growth to job mobility of this kind. However, individual workers also experience wage growth on a job. The following question is currently at the research frontier. Is observed on the job wage growth the consequence of training or is it to be explained by the workings of the matching process? This question is difficult to answer empirically because a formal test requires that one take account of sample selection induced by the matching process.

However, there are prior theoretical issues of empirical interest. For example, do the various versions of the investment in specific human capital hypothesis have different observable implications for the separation propensity and the probability distribution of the wage earned on the next job given separation or are they observationally equivalent, even in panel data? Given that the existence of differences in the wage paid across jobs for an individual and imperfect information about these differences are necessary for the very existence of job change in a stationary environment, can observations on the separation behavior and earnings of individuals over time provide any evidence concerning whether on the job training and/or job matching are important phenomena? These are some of the questions that motivate this paper.

The theoretical purpose of the paper is to formulate the on the job training and the job matching models of job to job movements within the same analytic framework, the search on the job model. Such a model facilitates analytic comparison of the implications of the special cases for separation or quit probabilities. Furthermore, a theoretical model of this kind is required to construct empirical tests that distinguishes between the two hypotheses using on the job wage observations.

Were information about the location and quality of worker-job matches perfect, turnover would not be observed in a stationary environment. Worker's would simply start and stay in the particular job offering the worker the highest lifetime wage, other job characteristics held constant. To the extent that the terms of employment on specific jobs are not known and time is required to obtain this information, the turnover process is one of moving from lower to high paying jobs along the lines of the Burdett's [1978] search model. How this process is modified when job specific productivity rises with tenure and when information about the job specific productivity improves with tenure is the focus of the analysis presented in the paper.

In its simplest form, Burdett's model can be characterized as follows. The wage paid is the worker's marginal productivity on the job. A worker's productivity on a particular job, given the worker's education, previous training, and other fixed indicators of general human capital, is regarded as a random draw from some distribution. Four additional assumptions characterize the model. First, productivity realizations across jobs for a particular worker are independent and identically distributed. Second, while on any job, information about the productivity on alternative jobs (offers) arrives infrequently. Third, productivity on each job, once realized, is stationary. Fourth, realized productivity is known with certainty once a job commences. The first assumption implies that productivity variation across jobs for an individual is match specific. The second captures the idea that time is required in the search process. The third and fourth assumptions are simplifications that are relaxed in the sequel.

Given all four assumptions, the model implies that workers move from higher to lower paying jobs whenever the opportunity arises and that the probability of separating from a job currently held is a decreasing function of the worker's wage and is independent of tenure on the job given the wage. The model can explain wage growth with general work experience and is consistent with a positive wage-tenure relation in cross section data as already noted. However, the simple search on the job model predicts no growth in earnings on a specific job and a probability of taking a wage cut equal to zero.

Mincer [1974] has long argued that observed wage growth on a specific job can be attributed to on the job training and/or learning on the job. This phenomena can be added to the model by an appropriate modification of the third assumption. The generalized model has the following implications. The probability of separation decreases with the current wage, given tenure as in the simple job search model. However, conditional on the wage earned and on measures of general human capital and the opportunity cost of time spent working, both the probability of a transition to a new job and the probability of a transition to non-employment increase (decrease) with tenure if the rate of job specific wage growth diminishes (increases) with tenure. Hence, if individual wage-tenure profiles are convex, as suggested by the evidence and economic principles, this result implies that the conditional distribution of completed job spell lengths given the wage should exhibit <u>positive</u> duration dependence. In addition, the probability that a worker will experience an initial reduction in wage when making a job to job transition is positive and increases with tenure. Both inferences are consequences of the fact that an alternative job currently paying the same wage offers more rapid future wage growth given the assumption that growth diminishes with tenure.

These implications of the on the job training model are either not drawn or are not fully appreciated in the existing literature on job turnover. Indeed, many authors claim that the on the job training hypothesis, explains why observed quit rates decline with tenure. This claim is valid only to the extent that it refers to the unconditional quit rate and only then because the quit rate is decreasing in the wage and because other factors that contribute to variation in quit rates across individual workers are not held constant. In other words, the theory implies that the observed negative duration dependence in the unconditional quit rate is the consequence of selectivity induced by the quit decision and unobserved heterogeneity.

Jovanovic [1979a, 1979b] has shown that learning about productivity on the job provides an alternative explanation for observed job specific wage growth. A pure form of this model is obtained by retaining the assumption of stationary productivity on a job but assuming that its "true" value must be inferred from the sequence of observations on the realized productivities that a worker experiences. Again quit rates fall with the wage earned given the individual's tenure, general human capital, and opportunity cost of time spent working. The job matching model also implies that the probability of quitting to take another job and the probability of taking a wage cut in the transition to another increase with tenure, given the current wage and measures of fixed individual differences. However, these implications follow as a consequence of the fact that an estimate of "true" match specific productivity becomes increasing more precise with experience. Specifically, because higher wage growth is more likely on a new job and because the worker always has the option to quit if an alternative proves to be a bad match which insures against the greater likelihood of lower growth, workers prefer jobs where individual productivity is more uncertain given the same current wage. Since the certainty about the value of a match increases with tenure, a worker is willing to accept a lower wage on an alternative job, given the wage earned on his or her current job, as tenure on the latter increases.

In contrast with the implications of the on the job training model, the argument outlined above does not imply a determinant sign for the conditional relationship between the probability of quitting to non-employment and tenure in general. However, one can establish that negative conditional duration dependence in the quit rate to non-employment is implied for all sufficiently large tenure values for those who choose to remain employed on a given job. In particular, if learning is complete in the limit as tenure increases, then those worker's whose wage exceeds the opportunity cost of working never quit to non-employment in the limit. Finally, the learning about the job hypothesis implies that the probability of a transition to non-employment increases with the opportunity cost of time spent working while this relationship is generally ambiguous in the on the job training case.

Recently Jovanovic [1984] has formulated a model similar to that of this paper which also combines his original matching model with a search on the job model and has obtained many of the same results as those reported here. However, the hypotheses used here to obtain implications are more general in at least two respects. First Jovanovic's approach requires unwarranted normality assumptions in the specification of the job specific wage process. These assumptions are not necessary in the formulation developed in this paper. Second, in order to establish that the separation decision is well defined, Jovanovic restricts parameter values of the model. Using the model developed in this paper, existence is established generally with no restrictions on parameter values other than natural non-negativity conditions. Finally, the model developed here is more useful for the purpose of empirical specification.

Given the similarity of the qualitative implications of the on the job training and the job matching models for separation rates, empirical studies that focus solely on turnover behavior cannot provide strong evidence that allows one to discriminate between them. Of course, non-positive conditional duration dependence in the job to job transition rate and a non-negative relation between the probability of taking a wage cut and tenure when making such a transition contradicts both hypotheses. However, it is well known that unobserved heterogeneity will bias downward the estimate of the effect of tenure on the probability of quitting to take another job. What seems to be required to test either hypothesis and/or to discriminate between them is a structured empirical formulation that simultaneously accounts for the observed wage dynamic and mobility of individuals that adequately accounts for unobserved fixed effects. The theoretical models developed in the paper will hopefully be useful for this purpose in future empirical research.

The text of the paper is presented in five major sections and two appendices. In the first section, the structure of the simple on the job search model is presented and its implications reviewed. Section 2 introduces a general model of separation decisions taken by an individual which includes both on the job training and job matching as special cases. The existence of a well defined separation decision problem under conditions that are far more general than those found in the literature is established in Appendix A. In section 3, the negative relationship between an individual's probability of both quitting to take another job and quitting to non-employment and the wage on the worker's current job conditional on tenure is shown to hold in the general model. The principal focus of the paper, the conditional quit-tenure relationship is the subject of section 4. Appendix B considers some specific examples of the Bayesian learning about the job model which serve to justify the general characterization given in section 4. The fifth section of the paper deals with the implications of the general model for empirical specification and estimation. The purpose of the section is to discuss methods that might be used to test for the presence of phenomena predicted by the on the job training and job matching hypotheses.

1. Search on the Job

In Burdett's [1978] model, a worker's wage on any job equals the marginal value of productivity, productivity on the job is stationary and certain, productivities are independently distributed across jobs, and the productivity realized on a job is known when the job commences. Let $F(x,m)$ denote the given cross job wage offer distribution with mean m and support W, a subset of the non-negative real numbers. Wage offers, independent random draws from the distribution, are assumed to arrive sequentially via a Poisson process with mean arrival rate $\lambda > 0$, which is the same whether employed or not. Workers move to an alternative job when an offer arrives if and only if the expected stream of discounted future earnings associated with the alternative is larger. Finally, workers are assumed to live forever, expectations are rational, and search is costless.

Under these assumption, the maximal expected discounted future stream of earnings is a stationary function of the wage currently earned and the mean of the wage offer distribution denoted as $v(w,m)$. The mean wage offer is included as a source of stationary worker heterogeneity in the model. Let t denote the length of the future time until the arrival of the next offer. Since t is distributed exponential with mean $1/\lambda$ and the offer x is distributed $F(x,m)$ given an arrival,

$$(1)\quad v(w,m) = E\{[(w/r)\ (1-e^{-rt}) + \max[v(x,m),\ v(w,m)]e^{-rt},$$

$$= \int_0^\infty [(w/r)\ (1-e^{-rt})\lambda e^{-\lambda t} dt$$

$$+ \int_0^\infty \int_W \max[v(x,m),\ v(w,m)]dF(x,m)\lambda e^{-(r+\lambda)t} dt$$

$$= w/(r+\lambda) + [\lambda/(r+\lambda)]\int_W \max[v(x,m),v(w,m)]dF(x,m)$$

where $r > 0$ is the interest rate. The equation states that the worker's expected wealth is equal to the expected present value of the current wage earned during the time until the next offer arrives plus the expected present value of expected wealth given the arrival of an alternative job offer given that the worker chooses among options optimally..

Obviously, the value function is only implicitly defined by (1), a functional equation. Notice that the value function is a fixed point of a transformation M, which maps the set of functions defined on the support of the wage offer distribution into itself. The transformed function is defined by

$$(2) \quad Mv(w,m) = w/(r+\lambda) + [\lambda/r+\lambda)]\int_W \max[v(x,m),v(w,m)]dF(x,m).$$

The existence of a unique fixed point, $v = Mv$, is established in Appendix A under assumption that the mean wage offer, m, is finite.

Given existence and uniqueness, it is the properties of the value function that are of interest for an analysis of turnover behavior. By showing that the transformation M maps the set of functions that are continuous and increasing in w into the set of strictly increasing continuous functions, one established that the unique fixed point is in the latter set. Given $v(w,m)$ continuous and increasing, it is obvious that (2) implies that the function $Mv(\cdot,m):W \to R$ is continuous and strictly increasing in w since each component on the right hand side has the property under the hypothesis. This method of proof is used throughout the paper.[1]

The fact that the value of a job is continuous and strictly increasing in the current wage earned implies that the set of acceptable alternative job offers is simply those offering a higher wage. This property of the model is sufficient to characterize the probability of a transition to another job and the distribution of the subsequent wage given such a transition. Specifically, the quit rate is equal to the product of the offer arrival rate and the probability that an offer exceed the worker's current wage,

$$(3.a) \qquad q(w,m) = \lambda[1-F(w,m)],$$

and the distribution of the worker's subsequent wage given a quit is the distribution of offers truncated by the worker's current wage,

$$(3.b) \qquad Q(x;w,m) = [F(x,m)-F(w,m)]/[1-F(w,m)], \; x > w.$$

Obviously, the quit rate is a decreasing function of the worker's current wage and the distribution of an acceptable alternative wage offer, the worker's wage on the next job, is stochastically increasing in w by virtue of (3).

In order to use the theoretical structure derived here and in the subsequent sections of the paper for empirical work, other features of the value function will be of interest. For example, it is useful to realize that the form of the value function depends on the form of the wage offer distribution. To make this point, first note that $v(w,m)$ is differentiable in w for fixed m and the derivative is

$$(4) \quad v'(w,m) = 1/(r+\lambda[1-F(w,m)])$$

by virtue of (1) if the offer c.d.f. is continuous. Hence, given a particular distribution function $F(w,m)$ one can solve for the form of $v(w,m)$ by integration in principle. Formally, (5.a) $$v(w,m) = c + \int\frac{dw}{r+\lambda[1-F(w,m)]}$$

where c is the constant of integration. Equation (1) also implies

(5.b) $$\lim_{w\to\bar{w}} \{rv(w,m)/w\} = 1, \text{ where } \bar{w} = \sup_{W} W,$$

is the least upper bound of the support of $F(\cdot,m)$. (5.b) is the boundary condition needed to determine the constant of integration, c. Some specific examples follow. In both examples the wage offer distribution has a single parameter m, which by construction is the mean.

Example 1. The Uniform Distribution: Let $F(w,m) = w/2m$, $w \varepsilon [0,2m]$, denote the wage offer distribution. Equation (5.a) implies

$$v(\cdot) = c + \int\frac{dw}{r+\lambda[1-w/2m]} = c - (2m/\lambda) \int\frac{(-\lambda/2m)dw}{r+\lambda[1-w/2m]}$$
$$= c - (2m/\lambda)\ln(r+\lambda)[1-w/2m])$$

Since $\bar{w} = 2m$, $v(\bar{w},m) = 2m/r$ by virtue of equation (5.b). Hence the constant of integration is $c = 2m/r + (2m/\lambda)\ln r$ and, consequently,

(6) $$v(w,m) = 2m/r + (2m/\lambda)\ln r - (2m/\lambda)\ln(r+\lambda[1-w/2m])$$
$$= (2m/r)[1 - (r/\lambda)\ln(1+(\lambda[1-w/2m]/r)].$$

Example 2. The Exponential Distribution: In this case $F(w,m) = 1 - \exp(-w/m)$, $w \varepsilon [0,\infty]$. Consequently, (5.a) implies

$$v(\cdot) = c + \int\frac{dw}{r+\lambda\exp(-w/m)} = c + (m/r) \int\frac{dy}{y+\lambda}$$

where $y = r\exp(w/m)$. Hence,

$$v(w,m) = c + (m/r)\ln(r\exp(w/m) + \lambda).$$

To satisfy (5.b), $c = -(m/r)\ln r$ is required which yields

$$(7)\quad v(w,m) = (m/r)\ln(\exp(w/m) + \lambda/r).$$

Of course, in both cases the expected wealth is an increasing function $v(w,m)$, of both the current wage and the mean of the wage offer distribution. But, note in addition that $v(w,m) = mv(w/m,1)$ i.e., the value function is homogeneous of degree one. Equivalently, a given percentage increase in the current wage and the mean of the wage offer distribution yields the same percentage increase in a worker's wealth. Of course, this result is not general. By virtue of Euler's theorem and equation (4), it holds if and only if the wage offer c.d.f. $F(x,m)$ is homogeneous of degree zero, as is the case in the two examples and many others as well.

This particular relationship between the form of the wage offer distribution and the value function is of interest for two reasons. First, the mean of the wage offer distribution is a measure of a worker's general ability across jobs and cross job relative variation of wage rates around the mean can be viewed as the dispersion of job specific ability when it is independent of the mean. Second, the assumption that $F(x,m)$ is homogeneous of degree zero yields specification restrictions on behavioral relations that are both useful for empirical work and testable. For example, equation (3) implies that the quit rate is homogeneous of degree zero in the wage and mean offer pair (w,m) i.e., $q(w,m) = q(w/m,1)$ if relative wage offer dispersion is independent of the mean in the sense that F is homogeneous of degree zero. Consequently, more able workers quit more frequently holding the current wage constant. An important empirical implication is that differences in general human capital must be corrected for in any attempt to estimate the effect of the current wage on the quit rate using cross section data since the model clearly implies that w and m are positively correlated in any data set. In subsequent sections, the implications of the specification are generalized.

2. A General Model of Job Turnover

Mincer [1974] argues that the observed positive cross section relationship between the wage earned on a job and tenure is simply the consequence of the accumulation of job specific skills through on the job training or learning. Jovanovic's [1979, 1984] formulation of the "job shopping" or matching hypothesis offers both an alternative and a complementary explanation. A worker's productivity on a specific job is uncertain but any estimate becomes more precise as more evidence is accumulated. Wage and tenure are positively associated because workers tend to remain on a job when the match is perceived to be a good one and because perceptions depend on wage experience. The principal purpose of this and subsequent sections of the paper is to develop and analyze a general extension of the basic search on the job model of turnover that includes these hypotheses as special cases. In this section, the structure of the general model is outlined.

As in the previous section, alternative offers arrive at an average given Poisson frequency equal to $\lambda > 0$ and are random draws from the cross job distribution of initial offers denoted as

(8) $\Pr\{w_0 \leq x\} = F_0(x,m)$

where m is the mean of the initial wage offer distribution, assumed to be finite, positive, and stationary for a given worker. It may be regarded as a measure of the worker's general human capital or earnings ability in the labor market.

In general, a worker will experience a sequence of stochastic wage rates on a specific job, so long as he or she remains on the job, denoted as $\{w_n\}$. The conditional distribution of the next wage in the sequence given that the current wage is w and the current sequence number is n follows:

(9) $\Pr\{w_{n+1} \leq x \mid w_n = w\} = F_{n+1}(x,w), \ (w,n) \in W \times N$

where W, a subset of the non-negative reals, is the set of possible wage rates and $N = \{0,1,...\}$ is the set of possible realizations of the job specific wage process.

Finally, the number of wage arrivals on a specific job to date is assumed to be generated by a Poisson process with arrival rate $\eta > 0$. Hence, if n(t) is the number of arrivals as of tenure date t,

(10) $\Pr\{n(t) = n\} = (\eta t)^n e^{-\eta t}/n!$

which of course has mean ηt. The time required until the next arrival of the job specific wage process is exponentially distributed, therefore, with mean $1/\eta$.

Since the number of realizations of the job specific wage process and tenure are positive correlated, n is referred to as tenure in the sequel. As further justification for this language, note that one can easily convert any monotone function of n into a corresponding monotone function of t by taking the expectation of the function with respect of the distribution given in (10). In other words, observed length of service on a job is an appropriate empirical surrogate for the number of realizations of a job specific wage process.

To complete the specification, let $b > 0$ denote a worker's stationary and non-stochastic value of time spent not employed, the non-employment benefit. Finally, the worker is assumed to act so as to maximize expected discounted streams of future income, where b is the income per period enjoyed when not employed and the wage is the income obtained when employed. The discount rate is $r > 0$ and the worker lives forever.

Let $u(b,m)$ denote the expected present value of future income when not employed given the mean of the worker's initial wage offer distribution m and the worker's non-employment benefit b. Given this number $u(b,m)$, the mean initial wage offer m, and any employment state pair (w,n), the expected present value of the worker's future income is a function $v_n(w,u,m)$. The number u and value function $v(\cdot,u,m):WxN \to R$ are derived using Bellman's principle as follows. Because the only future event that can change a worker's state when not-empolyed is the arrival of an offer, because the time to arrival is distributed exponential with mean $1/\lambda$, because the initial offer x given arrival of an offer is distributed $F_0(x,m)$, and because the worker accepts the offer if and only if its value is at least as large as that of non-employment, the value of non-employment satisfies the equation,

$$u(b,m) = E\{(b/r)(1-e^{-rt}) + \max[v_0(x,u(b,m),m),u(b,m))]e^{-rt}\} \tag{11}$$

$$= \int_0^\infty (b/r)\ (1-e^{-rt})\lambda e^{-\lambda t}dt$$

$$+ \int_0^\infty \int_W \max[v_0(x,u(b,m),m),u(b,m)]dF_0(x,m)\lambda e^{-(r+\lambda)t}dt$$

$$= b/(r+\lambda) + [\lambda/(r+\lambda)]\int_W \max[v_0(x,u(b,m),m),u(b,m))]dF_0(x,m).$$

where here t represent the random time required for the arrival of an offer. The first term represents the expected present value of a constant income stream equal to the non-employment benefit received over a future period equal to the time required for the arrival of the next offer and the second term is expected present value of the worker's optimal choice given the arrival of the next offer.

Given employment at a wage w after attaining tenure n in a specific job, the worker chooses between continued employment on the current job and quitting to take an alternative when an alternative offer arrives and chooses between continued employment and quitting to non-employment when a new job specific wage arrives. In either case, the quit option is exercised if and only if the value of the alternative is at least as large as the value of continuing on the job conditional on current information. Because the arrival of either one of these two events is Poisson with arrival rate equal to the sum $\lambda + \eta$ and because the conditional probability that an arrival will be an alternative job offer is $\lambda/(\lambda+\eta)$, the value of employment given the current wage tenure pair (w,n), value of non-employment u, and mean initial wage offer m satisfies the functional equation,

$$(12) \qquad v_n(w,u,m) = E\{(w/r)(1-e^{-rt})$$

$$+ \delta\max[v_0(x,u,m),v_n(w,u,m)]e^{-rt} + (1-\delta)\max[u,v_{n+1}(x,u,m)e^{-rt}\}$$

$$= \int_0^\infty (w/r)(1-e^{-rt})(\lambda+\eta)e^{-(\lambda+\eta)t}dt$$

$$+ \beta_1\int_0^\infty\int_W \max[v_0(\chi,u,m)v_n(w,u,m)]dF_0(x,m)(\lambda+\eta)e^{-(r+\lambda+\eta)t}dt$$

$$+ \beta_2\int_0^\infty\int_W \max[u,v_{n+1}(x,u,m)]dF_{n+1}(x,w)(\lambda+\eta)e^{-(r+\lambda+\eta)t}dt$$

$$= (1-\beta_1 - \beta_2)w/r$$

$$+ \beta_1\int_W \max[v_0(x,u,m),v_n(w,u,m)]dF_0(x,m)$$

$$+ \beta_2\int_W \max[u,v_{n+1}(x,u,m)]dF_{n+1}(x,w)$$

where t is the random time until arrival, $\delta = 1(0)$ if the arrival is an alternative offer (a new job specific wage), and

$$(13.a) \qquad \beta_1 = \lambda/(r+\lambda+\eta)$$

$$(13.b) \qquad \beta_2 = \eta/(r+\lambda+\eta).$$

The three terms on the right side of (12) respectively represent the expected present value of earnings until the next arrival, the product of the expected present value of the best choice given that the next arrival is an alternative wage offer and the probability of the specified event, and the product of the expected present value of the best choice given that the next arrival is a new wage on the job and the probability of the specified event.

Given the mean of the initial wage offer distribution m and any value of non-employment u, equation (12) implies that the value of employment function is a fixed point of the transformation T from the set of function defined on the space WxN to itself. The transformation T of any function $v(\cdot,u,m):W\times N \to R$ is defined by

$$
\begin{aligned}
(14) \qquad Tv_n(w,u,m) &= (1-\beta_1 - \beta_2)w/r \\
&+ \beta_1 \int_W \max[v_0(x,u,m), v_n(w,u,m)]\, dF_0(x,m) \\
&+ \beta_2 \int_W \max[u, v_{n+1}(x,u,m)]\, dF_{n+1}(x,w).
\end{aligned}
$$

In the remainder of the text the existence of a unique fixed point, $v = Tv$, for every finite pair (b,m) is assumed. In Appendix A, reasonable sufficient conditions imposed on the sequence of job specific wage rate distributions are shown to imply the assumption.[2]

3. The Reservation Offer and Wage

In this section the value of employment and value of non-employment functions are characterized and the implications of their properties for turnover behavior are derived. One assumption other than the existence of a unique value of employment is imposed. The on the job training, the job matching, and virtually every other conceivable hypothesis implies that the sequence of job specific wage rates are positively correlated. The natural formulation of this hypothesis is that the distribution of the next wage is stochastically increasing in the current wage, i.e., the next wage in the sequence is more likely to be larger the larger is the current wage. Formally, the hypothesis is that the distribution function $F_n(x,w)$ decreases in w for all n. Given this hypothesis, the value of employment is increasing in the wage currently earned.

Proposition 1: The value of employment function $v_n(w,u,m)$ is continuous and strictly increasing in w if $F_n(x,w)$ is continuous and decreasing in w for all n.

Proof. Because the map T is a contraction, it satisfies Blackwell's [1965] sufficient conditions and because the set of continuous increasing functions is a metric space, T has a unique fixed point. Hence, it is sufficient to show that T maps the set of functions with the asserted properties into itself.

Since max[y,z] is continuous in (y,z), equation (14) implies that Tv is continuous in w if v and F are continuous in w. Suppose that v is non-decreasing in w. Then the first term on the right of (14) is strictly increasing, the second term is non-decreasing by virtue of the fact that the function max[y,z] is increasing in z, and the last term is non-decreasing by virtue of the fact that the expectation of a non-decreasing function of a random variable which is stochastically increasing in w is non-decreasing.

Because a unique value of employment exists for every finite value of non-employment and the right side of equation (12) is continuous in both u and v, the value of employment is also a continuous function of u for all fixed (w,n) by virtue of the implicit function theorem. Not surprisingly, the value of employment is increasing in the value of non-employment but a given increase in the latter induces a lesser increase in the former.

Proposition 2: The value of employment function $v_n(w,u,m)$ is continuous and increasing in u. Furthermore, $v_n(w,u+c,m) \leq v_n(w,u,m) + (\beta_1 + \beta_2)c < v_n(w,u,m) + c$ for every constant c and fixed $(w,n) \in W \times N$.

Proof: Equation (14) obviously implies that T maps the set of functions continuous and increasing in u into itself and hence its unique fixed point must have this property. Consider any function with the property that $v_n(w,u+c,m) < v_n(w,u,m) + c$ for all $(w,n) \in W \times N$. By virtue of (14) and the fact that (13) implies $\beta_1 + \beta_2 < 1$,

$$Tv_n(w,u+c,m) = w/(r+\lambda+\eta)$$

$$+ \beta_1 \int_W \max[v_0(\chi,u+c,m),v_n(w,u+c,m)]dF_0(\chi,m)$$

$$+ \beta_2 \int_W \max[u+c,v_{n+1}(\chi,u+c,m)]dF_{n+1}(x,w).$$

$$< w/(r+\lambda+\eta) + \beta_1 \int_W \max[v_0(x,u,m),v_n(w,u,m)]dF_0(x,m)$$

$$+ \beta_2 \int_W \max[u,v_{n+1}(x,u,m)]dF_{n+1}(x,w) + (\beta_1 + \beta_2)c$$

$$= Tv_n(w,u,m) + (\beta_1 + \beta_2)c < Tv_n(w,u,m) + c.$$

Because T maps the set of functions with the property into itself its unique fixed point has the property.

The fact that the value of non-employment exists and is unique for every finite benefit-mean pair (b,m) and is an increasing function of b is an important implication of Proposition 2.

Proposition 3: A unique positive value of non-employment exists for every finite positive benefit-mean pair (b,m) and the function u(b,m) is continuous and strictly increasing in b.

Proof. Given (b,m) finite, u(b,m) solves $f(u,b,m) = 0$ where $f(u,b,m) = b/(r+\lambda) + [\lambda/(r+\lambda)]\int_W \max[v_0(x,u,m),u]dF_0(x,m) - u$. by virtue of equation (11). Given $b > 0$, it is clear that $\infty > f(0,b,m) \geq b/(r+\lambda) > 0$. Proposition 2 implies that f(u,b,m) is continuous and strictly decreasing in u. Furthermore,

$$
\begin{aligned}
f(c,b,m) &= b/(r+\lambda) + [\lambda/(r+\lambda)]\int_W \max[v_0(x,c,m),c]dF_0(x,m) - c \\
&\leq b/(r+\lambda) + [\lambda/(r+\lambda)]\int_W \max[v_0(x,0,m)+c,c]dF_0(x,m) - c \\
&\leq f(0,b,m) - [r/(r+\lambda)]c < 0
\end{aligned}
$$

for all $c > (r+\lambda)f(0,b,m)/r$. Hence, $f(u,b,m) = 0$ has a unique, positive, and finite solution by virtue of the mean value theorem.

Because f(u,b,m) is continuous and strictly decreasing in u and continuous and strictly increasing in b, the function u(b,m) defined by $f(u(b,m),b,m) = 0$ is continuous and strictly increasing in b by virtue of the implict function theorem.

Given Proposition 3, the value of employment can be expressed as a function of the worker's state, the pair (w,n), and the benefit-mean pair (b,m). Specifically, $v_n(w,u(b,m),m)$ is a continuous and increasing function of both the wage and the non-employment benefit by virtue of Propositions 1-3. Finally, it will be useful to show that a worker always prefers employment at a wage equal to the non-employment benefit.

<u>Proposition 4</u>: For all $n \in N$ and fixed finite (b,m)

$$u(b,m) < v_n(b,u(b,m),m).$$

Proof: Given (b,m) equation (11) can be rewritten as

$$u = (1-\beta_1-\beta_2)b/r + \beta_1\int_W \max[v_0(x,u,m),u]dF_0(x,m) + \beta_2 u$$

by virtue of (13). Hence, if $v_n(b,u,m) \geq u$ for all n, then (14) implies

$$\begin{aligned} Tv_n(b,u,m) &= (1-\beta_1 - \beta_2)b/r \\ &+ \beta_1\int_W \max[v_0(\chi,u,m),v_n(b,u,m)]dF_0(x,m) \\ &+ \beta_2\int_W \max[u,v_{n+1}(x,u,m)]dF_{n+1}(x,b). \\ &> (1-\beta_1-\beta_2)b/r + \beta_1\int_W \max[v_0(x,u,m),u]dF_0(x,m) + \beta_2 u = u. \end{aligned}$$

Continuity and strict monotonicity of the value of employment function in w implies that the decisions to accept a wage offer when not employed, to move to another job when an opportunity arises, and to quit to non-employment when conditions on the job change all have the reservation property. Specifically, a reservation offer exists which defines the lower bound on the set of acceptable alternative wage offers and a reservation wage exists which defines the lower bound on the set of job specific wage rates that are preferred to non-employment given the worker's employment state. The reservation wage offer given employment characterized by the wage-tenure pair (w,n) is a function $\rho_n(w,b,m)$ implicitly defined as the solution to

$$(15) \qquad v_0(\rho_n(w,b,m),u(b,m),m) = v_n(w,u(b,m),m).$$

An employed worker in state (w,n) will accept any alternative job with wage offer in excess of $\rho_n(w,b,m)$. The arrival of a new wage on the job less than the reservation wage, a function $\sigma_n(b,m)$ that solves

$$(16) \qquad v_n(\sigma_n(b,m),u(b,m),m) = u(b,m),$$

will induce the worker to quit to non-employment. Finally, the smallest acceptable wage offer when not employed is simply $\sigma_0(b,m)$.

Formally, Proposition 1-4 have the following implications:

Corollary 1: The reservation offer $\rho_n(w,b,m)$ is a continuous and strictly increasing function of w for all n. The reservation wage $\sigma_n(b,m)$ is a continuous and strictly increasing function of b for all n. Furthermore $\sigma_n(b,m) \leq b$.

Proof. The first statement is implied by Proposition 1, the second is implied by Propositions 2 and 3, and the final statement is implied by Propositions 1 and 4.

Obviously, there are two kinds of quit behavior predicted by the general model, quits to a new job and quits to non-employment. The associated quit rates are technically the probabilistic rates of transition from one to another job and from employment to non-employment respectively. Both depend on the worker's current employment state, the wage-tenure pair, the non-employment benefit, and the mean of the initial wage offer distibution. The transition rate to another job is the product of the rate at which offers arrive and the probability that employment at the alternative job given no tenure is preferred to continued employment in the current job, i.e., it is the function

$$(17) \qquad q_n(w,b,m) = \lambda[1-F_0(\rho_n(w,b,m),m)].$$

Analogously, the transition rate to non-employment is equal to the product of the rate at which new job specific wage rates arrive and the probability that the next arrival will be less than the reservation wage, i.e., the function

$$(18) \qquad p_n(w,b,m) = \eta F_{n+1}(\sigma_{n+1}(b,m),w).$$

Finally, a worker's rate of transition from non-employment to acceptable employment is the product of the offer arrival rate and the probability that an offer is acceptable,

$$(19) \qquad \alpha(b,m) = \lambda[1-F(\sigma_0(b,m),m)].$$

Given equations (15) - (19), Corollary 1 implies that both the job to job transition rate and the job to non-employment transition rate are deceasing functions of the current wage. These implications are consistent with the general empirical observation that quit probabilities are negatively associated with the wage earned at the time of the separation. However, it is important to note that this inference of the model in cross section data is conditional on heterogeneity across worker's reflected in both the mean initial wage offer and the non-employment benefit, our measures of general human capital and opportunity cost of time spent working respectively. Of course, Corollary 1 also implies that the job to non-employment quit rate is increasing in the non-employment benefit and that the rate of transition from non-employment is a decreasing function of b.

Although the expected present value of future income is increasing in the non-employment benefit at all tenures by virtue of Propositions 2 and 3, the direction of the effect of an increase on the reservation wage depends in general on the specification of the distributions of the sequence of future on job specific wage rates. Specifically, equation (15) implies that the qualitative effect of an increases in b on the reservation offer depends on the relative sizes of the effects of an increase in the value of non-employment on the values of employment at different tenures. It not clear that the net effect can be signed without additional assumptions.

In the general model, the wage a worker is observed to earn changes from time to time either because the worker moves to a new job or because the wage on the current job is revised. The distributions of the new wage in both cases depend on the worker's employment state as a consequence of the fact that a worker's decision to either accept the alternative or to continue on the job depends on the state. The distribution of the new wage given a job to job transition is the conditional initial wage offer distribution given that the offer exceeds the reservation offer. Let

$$(20)\quad Q_n(x;w,b,m) = \begin{cases} 0 \text{ if } x < \rho_n(w,b,m) \\ \dfrac{F_0(x,m) - F_0(\rho_n(w,b,m),m}{1-F_0(\rho_n(w,b,m)} \text{ otherwise} \end{cases}$$

denote the distribution of the new wage given a job to job transition. Corollary 1 implies that the wage distribution on an alternative job is stochastically increasing in w, i.e., the wage earned on a subsequent job is more likely to be higher the higher is the current wage. Note that this result does not require positive expected growth in the sequence of job offers. It is the consequence of selection induced by the worker's turnover decisions. Finally, the conditional distribution of the next job specific wage given that the worker stays on the job is the unconditional distribution truncated below by the reservation wage, i.e.,

$$(21)\quad P_n(\chi;w,b,m) = \begin{cases} 0 \text{ if } x < \sigma_{n+1}(b,m) \\ \dfrac{F_{n+1}(x,w)-F_{n+1}(\sigma_{n+1}(b,m),w))}{1-F_{n+1}(\sigma_{n+1}(b,m),w)} \text{ otherwise.} \end{cases}$$

Earlier, we argued that it was natural to assume wage offer dispersion which is independent of the mean in the sense that $F(x,m) = F(x/m,1)$ or equivalently $x = ym$ where y is distributed independent of m with unit mean.[3] This restriction applied to the distribution of subsequent wage rates is also consistent with both the job matching and the on the job training hypotheses. These restrictions imply that both the value of employ- nt function and the value of non-employment function are homogeneous of degree one.

Proposition 5: The value functions are of the forms

$$(22) \qquad v_n(w,u(b,m),m) = mv_n(w/m,u(b/m,1),1)$$

and

$$(23) \qquad u(b,m) = mu(b/m,1)$$

if $F_0(x,m) = F_0(x|m,1)$ and $F_n(x,w) = f_n(x|w,1)$.

Proof: Consider a value of employment function of the form specified in (22). Under the hypothesis, equation (11) can be written as

$$(24) \qquad u(b,m)/m = (b/m)/(r+\lambda)$$

$$+ [\lambda/(r+\lambda)]\int \max[v_0(y,u(b/m,1),1),u(b,m)/m]dF_0(y,1)$$

which implies $u(b,m)/m = u(b/m,1)$. Hence, equation (14) implies that the transformation of any function of the form specified in (22) under the hypothesis is

$$(25) \qquad Tv_n(w,u(b,m),m) = (1-\beta_1-\beta_2)w/r$$

$$+ \beta_1\int \max[mv_0(y,u(b/m,1),1),\ mv_n(w/m,u(b/m,1),1)]dF_0(y,1)$$

$$+ \beta_2 \int \max[mu(b/m,1), mv_{n+1}(yw/m,u(b/m,1)]dF_{n+1}(y,1)$$

$$= mTv_n(w/m,u(b/m,1),1).$$

under the hypothesis. Consequently, the unique fixed point of the transformation is of the asserted form.

In other words, an equal proportional increase in w,b, and m increases wealth when employed by the same proportion and an equal proportional increase in b and m increases wealth when not employed by the same proportion. That both are increasing functions of the mean wage offer is an implication of the proposition.

One behavioral implication of Proposition 5 is that the quit rates and truncated wage distributions functions defined above depend only the ratios of the arguments.

Corollary 2: Under the hypothesis to Proposition 5, the reservation offer and reservation wage functions are both homogeneous of degree one. Formally:

(26) $\rho_n(w,b,m) = m\rho_n(w/m,b/m,1).$

(26) $\sigma_n(w,b,m) = m\sigma_n(w/m,b/m,1).$

The corollary together with equations (17)-(19) have the following implications. First, because the effect of the non-employment benefit on the job to job transition rate is ambiguous, the effect of an increase in mean wage offer is also ambiguous. Second, because the employment to non-employment transition is decreasing in the current wage and increasing in the non-employment benefit, the effect of an increase in the mean wage offer is again ambiguous. Finally, the effect of an increase in the mean wage offer on the non-employment to employment transition rate is positive because the transition rate is $\alpha(b/m,1)$ and because α in generally decreasing in b.

4. The Quit-Tenure Relation

The principal implications of simple search on the job model are generalized in the previous section. The distinctive features of both the on the job training and the job matching hypotheses is the dependence of the quit rates on tenure. The purpose of this section is to delineate the implications of these hypotheses for the quit-tenure relation.

The two hypotheses can be distinguished as follows: On the job training implies that the sequence of expected job specific wage increments are positive but diminish with tenure while job matching implies that the sequence of expected wage increments are all zero but their "variances" diminish with tenure. Formally,

$$(28.a) \qquad \int_W (x-w)dF_n(x,w) \geq 0$$

$$(28.b) \qquad F_n(x,w) \leq F_{n+1}(x,w) \text{ for all } x \in W$$

in the training on the job case. In other words, the job specific wage is expected to grow but the expected growth rate diminishes with tenure. In the job matching case, the mean of the next wage is the current wage for all tenure but the dispersion of the distribution of the next job specific wage, in the sense of the Rothschild-Stiglitz [1970] definition of mean preseving spread, decreases with tenure, i.e.,

(29.a) $\quad \int_W (x-w)dF_n)(F_n(x,w) = 0$

(29.b) $\quad \int_0^z F_n(x,w)dx > \int_0^z F_{n+1}(x,w)dx$ for all $z \varepsilon W$

in the job matching case.

Condition (28.b), that larger wage increments are more probable at lower tenures, captures the idea that most learning and other forms of specific capital accumulation occur earlier in the tenure on a job rather than later. The assumption is generally required to obtain the implication that the expected wage increment on a specific job given the current wage, the left side of (28.a), declines with tenure as the data suggest. In Jovanovic's learning about productivity on the job formulation, the wage paid at any data is the current Bayesian estimate of the worker's true productivity on the job. Condition (29.a) is an implication of that fact. The wage, i.e., estimate of job productivity, is revised from time to time in response to observations on realized productivity. As more observations are accumulated with tenure on the job, the wage become a more precise estimate of job specific productivity and as a consequence the dispersion of the distribution of the next wage given the current wage declines with tenure. The latter idea is captured here by interpreting n as the size of the sample of productivity observations to date and by assuming that variation in the next wage declines with sample size in the sense that the mean preserving spread of the next wage declines with n. In appendix B, two different specifications of a Bayesian learning model are developed and shown to imply (29).

The remainder of the section is devoted to establishing the principal result of the paper. Namely, both hypotheses as formalized in (28) and (29) imply that the value of employment conditional on the current wage, non-employment benefit, and mean wage offer declines with tenure. Consequently, both imply that the conditional reservation offer declines with tenure and the job to job transition rate increases, i.e., exhibits positive duration dependence. Although the principal result also implies that the on job reservation wage increases with tenure, the tenure effect on the job to non-employment transition rate increases with tenure given the on the job training hypothesis but is ambiguous in general in the case of the job matching hypothesis.

Proposition 6: The value of employment function $v_n(w,u,m)$ is decreasing in n if $F_n(x,w)$ is decreasing in w and increasing in n.

Proof: in the proof, the fixed argument (u,m) is implicit. Equation (14) and the hypothesis imply

$$Tv_n(w) - tv_{n+1}(w) =$$

$$\beta_1 \int \{\max[v_0(x),v_n(w)] - \max[v_0(x),v_{n+1}(w)]\} dF_0(x,m) +$$

$$\beta_2 \int \max[v_{n+1}(x),u] dF_{n+1}(x,w) - \beta_2 \int \max[v_{n+2}(x),u] dF_{n+2}(x,w)$$

$$= \beta_1 \{\max[v_0(x),v_n(w)] - \max[v_0(x),v_{n+1}(w)]\} dF_0(x,m +$$

$$\beta_2 \int \{\max[v_{n+1}(x),u] - \max[v_{n+1}(x),u]\} \; dF_{n+1}(x,w)$$

$$\beta_2 \int \max[v_{n+2}(x),u][dF_{n+1}(x,w) - dF_{n+2}(x,w] > 0$$

because the first and second terms are non-negative if v is decreasing in n and the last term is non-negative by virtue of the fact that the expectation of an increasing function, $v_{n+2}(x)$, of a random variable taken with respect to a distribution that stochastically dominates another exceeds the expectation taken with respect to the latter. Of course, the value function at every tenure is strictly increasing in the wage by virtue of Proposition 1 under the hypothesis. Hence, the contraction T maps the set of functions that are decreasing in n into itself and, consequently, its unique fixed point in decreasing in n.

Propositions 1 and 6 have the following impolication given equations (15) and (16):

Corollary 3: The reservation offer $\rho_n(w,b,m)$ is decreasing n, and the reservation wage $\sigma_n(b,m)$ is increasing in n under the hypothesis to Proposition 6.

Hence, both quit rates exhibit positive duration dependence conditional on the current wage earned by virtue of equations (17) and (18). The training on the job model implies that the conditional job to job transition rate increases with tenure when job specific wage increments diminish with tenure because an alternative job at the same wage is more likely to offer more rapid wage growth in the future. A positive tenure effect in the case of the conditional job to non-employment transition rate is implied for virtually the same reason. A positive tenure effect in the case of the conditional job to non-employment transition rate is implied for virtually the same reason. Given the wage, the worker can expect less wage growth in the future on a specific job as tenure increases. For the same reasons, the distribution of acceptable alternative wage offers is stochastically decreasing in n by virtue of equations (20). Consequently, a worker is more likely to accept a wage cut when moving to another job the more tenure he or she has accumulated on the current job. Although the later implication of the model is acknowledged in the existing literature, the associated implication that quit rates condition on the wage should exhibit positive duration dependence does not seem to be appreciated with the notable exception of Jovanovic [1984].

In the case of the job matching model, the value of employment declines with tenure if the worker "prefers risk" in the sense that the value function in convex in the wage as the following result demonstrates.

Proposition 7: The value function $v_n(w,m)$ decreasing in n if $v_n(w,m)$ is increasing and convex in w and $F_n(x,w)$ is a mean preserving spread of $F_{n+1}(x,w)$ for all n.

Proof: Equation (14) and the hypothesis imply

$$Tv_n(w)-Tv_{n+1}(w) = \beta_1\int\{\max[v_0(x),v_n(w)]-\max[v[x_0), v_{n+1}(w)]\}dF_0(x,m)$$

$$+ \beta_2[\int\max[v_{n+1}(x),u]dF_{n+1}(x,w) - \int\max[v_{n+2}(x),u]dF_{n+2}(x,w)]$$

$$= \beta_1\int\{\max[v_0(x), v_n(w)]-\max[v_0(x),v_{n+1}(w)]\}dF_0(x,m)$$

$$+ \beta_2\int\{\max[v_{n+1}(x),u] - \max[v_{n+2}(x)u]\}dF_{n+1}(x,w)$$

$$+ \beta_2[\int\max[v_{n+2}(x),u]dF_{n+1}(x,w) - \int\max[v_{n+2}(x),u]dF_{n+2}(x,w)].$$

The first and second terms are non-negative if $v_n(w,m) \geq v_{n+1}(w,m)$ for all $(w,n) \in WxN$. If $v_{n+2}(x,m)$ is convex in x, then the last term is non-negative because $F_{n+1}(x,m)$ is a mean preserving spread of $F_{n+2}(x,m)$ by assumption. Hence, $v_n(w,m) \geq v_{n+1}(w,m)$ for all $(w,n) \in WxN$ implies $Tv_n(w,m) \geq Tv_{n+1}(w,m)$. Since T is a contraction, the argument implies that its unique fixed point is decreasing in n under the hypothesis.

Preference for risk is a common property of stopping models because the decision maker generally has the option of rejecting unfavorable random realizations. For example, note that equation (14) immediately implies that the transformation T maps the set of convex functions into itself in the special case of $\eta = 0$, the simple search on the job model. The result holds in the general case as well if the distribution of the next proportional increment to the wage is independent of the current wage.

Proposition 8: If $F_n(x,w) = F_n(x/w,1)$ for all $n > 0$, $v_n(w,b,m)$ is convex in w for all n.

Proof. Under the hypothesis, (14) implies

$$Tv_n(w) = w/(r+\lambda+\eta) + \beta_1 \int_W \max[v_0(x), v_n(w)]\,dF_0(x,m)$$

$$+ \beta_2 \int_0^\infty \max[v_{n+1}(yw), u]\,dF_{n+1}(y,1)$$

The first term on the right is linear in w and the second term and third terms are convex functions of w if $v(\cdot)$ is convex of w since $\max[y,z]$ is convex in (y,z). Hence, T maps the set of convex functions into itself and consequently its unique fixed point is in the set.

<u>Corollary 4:</u> The reservation offer $\rho_n(w,b,m)$ is decreasing n, and the reservation wage $\sigma_n(b,m)$ is increasing in n if $F_n(x,w)$ is a mean preserving spread of $F_{n+1}(x,w)$ and $F_n(x,w) = F_n(x/w,1)$ or all n.

Hence, the job to job transition rate and the distribution of acceptable alternative wage offers have the same qualitative properties in both models by virtue of equations (17) and (20). The conditional transition rate to another job increases with tenure and the distribution of acceptable alternative wage offers is stochastically decreasing in tenure. However, the reason for the results in the job matching case is different. An untried alternative job that initially offers the same wage as that currently earned is preferred because the worker knows that the current job is less likely to offer higher wage rates in the future than an untried alternative and because the worker has the option to quit the alternative should it prove to me a bad match. The option serves to insure against down side risk and is the reason for the worker's derived "preference for risk" in the model.

However, Corollary 4 does not have the same implication for the relationship between the quit to non-employment rate and tenure in the job matching case. In fact the qualitative effect cannot be signed in general. An inspection of equation (18) reveals two sources of the ambiguity. First, the hypothesis that dispersion in the sense of mean preserving spread diminished with tenure is not sufficient to sign the effect of n on the value of the c.d.f. at w in general. Second, even in a case like the normal distribution where $F_{n+1}(x,w) > (<) F_n(x,w)$ as $x \geq (\leq) w$, an implication of the fact that the variance is one to one with the notion of mean preserving spread, the effect of tenure on the job to non-employment transition rate depends on whether the wage earned at tenure n is greater or less than the reservation wage at tenure n+1. For similar reasons, the effect of tenure on the distribution of the next acceptable job specific wage is also ambiguous.

However, if one assumes that learning about a worker's productivity on a specific job is complete in the limit as tenure becomes large, the following result implies that the conditional job to non-employment transition rate exhibits negative duration dependence given any wage greater than or equal to the non-employment benefit and positive duration dependence otherwise for all large tenure values.

Proposition 9: If $F_n(x,w)$ is a mean preserving spread of $F_{n+1}(x,w)$ and

(30) $$\lim_{n\to\infty} F_n(x,w) = \begin{cases} 0 \text{ if } x < w \\ 1 \text{ otherwise} \end{cases}$$

then

(31) $$\lim_{n\to\infty} p_n(w,b,m) = 0 \text{ if } w \geq b.$$

Proof. By virtue of equations (18) and (30) it is sufficient to show that the sequence $\{\sigma_n(b,m)\}$ limits to b as n tends to infinity. That the sequence has a limit is a consequence of the previously established facts that it is an increasing sequence bounded by b. Given the continuity of the value function in w, the limit of the sequence is b if the limit of $v_n(b,b,m)$ is $u(b,m)$ by virtue of equation (16). Finally, (30) and equations (12) and (8.b) imply

$$\lim_{n\to\infty} v_n(b,b,m) =$$

$$(b/r)/[1-\beta_1 - \beta_2] + \beta_1 \int \max[v_0(x,b,m), \lim_{n\to\infty} v_n(b,b,m)]dF_0(x,m)$$

$$+ \beta_2 \max[u(b,m), \lim_{n\to\infty} v_{n+1}(b,b,m)].$$

Hence, (31) is implied by (11) and (13).

Finally, since the proposition implies virtually all workers who attain long tenures will have a wage in excess of the non-employment benefit, i.e., the reservation wage $\sigma_n(b,m)$ is "close" to b for all large n, only negative duration dependence can be observed in any data set. In other words, the selection that results as a consequence of the decision to remain employed on a specific job implies that only workers with wage rates such that the conditional duration dependence is negative will be in any observed sample of workers with long tenure.

5. Stochastic Structure and Estimation with Panel Data

Given the general model developed in the paper, an individual's employment history can be modelled as a continuous time Markov process provided one defines the state space appropriately. A "state" is a specification of whether or not the worker is employed and the wage-tenure pair (w,n) given employment. The probability of finding an individual in a state at some future date given the worker's history is a function of only the worker's current state and the length of the time interval between the current and future dates. This fact provides the means for specifying the likelihood of individual job histories needed to test the empirical validity of the various job turnover models reviewed in this paper. The purpose of this section is to sketch a method for constructing appropriate likelihood functions.

One way of viewing a worker's employment history is as a time ordered sequence of job spells interrupted occasionally by spells of non-employment. One can construct likelihoods of these sequences using the following concepts. Given non-employment, the only possible transition in a infinitesimal time interval is to employment at an acceptable wage offer, a draw from the distribution of initial wage offers initial wage offers $F_0(\cdot)$ that is no less than the reservation wage σ_0. The instantaneous probability of an offer arrival given non-employment is λdt. Given employment characterized by the wage-tenure pair (w,n), one of three instantaneous transitions is possible. If a new job specific wage arrives, then a transition occurs which is either to employment at the new wage and tenure pair when the new wage is acceptable or to non-employment when it is not. The new wage is a random draw from the c.d.f. $F_{n+1}(\cdot,w)$ and is acceptable if and only if greater than or equal to the reservation wage σ_{n+1}. If an acceptable alternative wage offer arrrives, one drawn from the distribution $F_0(\cdot)$ that is no less than the reservation offer $\rho_n(w)$, then a transition to a new job, characterized by the wage offered and no tenure, is made. Finally, the instantaneous probability of a new job specific wage arrival is ηdt and of an alternative wage offer is λdt.

A history of a completed job is characterized by an ordered sequence of k wage-duration pairs $\{w_{i-1},t_i\}$, where t_i represents the length of time wage w_{i-1} was earned and $k \in \{1,2,\cdots\}$ is the total number of different wage rate experienced on the job, and an indicator of how the spell ends, $\delta \in \{0,1\}$ where $\delta=0$ signals that the spell ends with a direct transition to another job and $\delta=1$ signals a transition to non-employment. Let $\Pi(\{w_{i-1}, t_1\}, k, \delta, w_0)$ represent the conditional probability density of a specific job history given the initial wage. It can be constructed as follows.

Since a give wage subspell ends when either a new job wage arrives or an acceptable alternative offer arrives, the conditional distribution of the spell length given the state is exponentially distributed with constant hazard equal to the sum of the two arrival rates, $\eta+q_{i-1}(w_{i-1})$, where the latter is the conditional quit rate to an alternative job as specified in equation (17). Given the model, the fact that the job continues beyond spell i for $i < k$ implies that such a spell ended with the arrival of an acceptable new job specific wage. This event has conditional probability equal to the rate at which acceptable new job specific wage rates arrive given the state, $\eta[1-F_i(\sigma_i,w_{i-1})]$, divided by the hazard rate for the distribution of the duration of the subspell. Of course, the conditional probability density to associate with the next observed wage is $f_i(w_i,w_{i-1})/[1-F_i(\sigma_i,w_{i-1})]$, where $f_i(\cdot, w_{i-1})$ is the p.d.f. associated with the c.d.f. $F_i(\cdot,w_{i-1})$, since w_i was accepted. Finally, the last spell i=k ended with the arrival of either an acceptable alternative offer, signalled by $\delta=0$, or a new job specific wage which was not acceptable, indicated by $\delta=1$. The conditional probability of the former event is simply the conditional quit rate to an alternative $q_{k-1}(w_{k-1})$ divided by the hazard rate in the final subspell.

After appropriate cancellations are made, the discussion in the previous paragraph implies the following specification of the likelihood of a completed job spell:

$$\pi(\{w_{i-1},\ t_i\},m,\ \delta;w_0\} \quad =$$

$$\overset{k-1}{\underset{1}{\chi}} \exp\ \{-[\eta+q_{i-1}\ (w_{i-1})]t_i\}\eta f_i(w,\ w_{i-1})$$

$$\chi\ \exp\{-[\eta+q_{k-1}(w_{k-1})]t_k\}[q_{k-1}\ (w_{k-1})]^{1-\delta}[\eta F_k(\sigma_k,w_{k-1})]^{\delta}.$$

if $w_i > \sigma_i$ for $< k$. Equivalently,

$$\pi(\{w_{i-1},t_i\},k,\delta;w_0\} \ = \eta^{\delta}\eta^{k-1}\exp\{-\eta\sum_i^k t_i\}\ \chi \tag{32}$$

$$[q_{k-1}\{w_{k-1})]^{1-\delta}\ \overset{k}{\underset{i}{x}}\ \exp\{-q_{i-1}(w_{i-1})t_i\}\ \chi$$

$$F_k(\sigma_k,w_{k-1})]^{\delta}\ \overset{k-1}{\underset{1}{x}}\ \ f_i(w_i,w_{i-1})\ \text{if}\ w_i > \sigma_i\ \text{for all}\ i > k$$

and zero otherwise. The first term on the right is proportional to the probability that k-1 wage changes occur during an interval equal in length to the total job spell duration, the sum of the wage subspell durations, if the job spell ends with a quit to an alternative job (δ=0). The term is proportional to the probability of k wage changes if the job spell ends with a quit to non-employment because such a transition occurs instantaneously after the arrival of an unacceptable wage on the job. The second term is the probability that no quit to an alternative job takes place during the first k-1 wage spells of observed lengths times the probability density that the last spell has the duration observed if δ=0. If δ=1, then the second term is simply the probability that the worker does not quit to take another job during any of the recorded k-1 wage subspells since the spell ends with a quit to non-employment in this case. There are k wage changes but the value of the last is not observed when $\delta=1$. In this case, the last term is the joint probability density associated with the observed sequence of k-1 wage rates times the probability then the last is not observed as a consequence of a quit to non-employment, i.e., a realization less than the reservation wage. Finally, if $\delta=0$, the last term is simply the joint unconditional density of the observed sequence of wage rates.

The fact that the value functions associated with the decision problem and consequently the reservation wage rates and the reservation offers do not depend on the rate at which new job specific wage rates arrive is an important (implicit) implication of the general model. Consequently, η is a "nuisance" parameter for the problem of estimating the model; i.e., the first term on the right of (32) can be ignored in the specification of a likelihood function for any sample.

A completed job spell is followed by a transition to another job. Transitions from one job to another are of two types as we have already noted; either the worker's spell on the origin job ends with an immediate transition to another job or there is an intervening period of non-employment. Formally, let $\phi(w_0,t_0;w,n,\delta)$ denote the joint probability density assignable to the observation that the initial wage earned on the destination job is w_0 after an intervening period of length t_0 given that the worker's employment state was (w,n) on the origin job at the time of the transition and the transition was of type $\delta \in \{0,1\}$. Given $\delta = 0$, $t_0 = 0$ with probability 1 and w_0 is a random drawn from the distribution of acceptable initial wage offers given employment in state (w,n). Consequently,

$$(33) \qquad \phi(w_0,t_0;w,n,0) = \begin{cases} f_0(w_0)/[1-F_0(\rho_n(w\ ))] \text{ if } w_0 \geq \rho_n(w). \\ 0 \text{ otherwise.} \end{cases}$$

However, if the prior job spell ends with a transition to non-employment, then the intervening non-employment duration is exponential with hazard $\lambda[1-F_0(\sigma_0)]$ and w_0 is a random acceptable initial offer which has distribution $f_0(w_0)/[1-F(\sigma_0)]$. Hence, the joint probability density of the wage duration pair is

$$\phi(w_0,t_0;w,n,1) = \begin{cases} \lambda\exp\{-\lambda[1-F_0(\sigma_0)]t_0\}f_0(w_0) \text{ if } w_0 \geq \sigma_0. \\ 0 \text{ otherwise.} \end{cases} \tag{34}$$

As both the on the job training and the matching hypotheses imply that the reservation wage increases with tenure and reservation offer decreases with tenure, as measured by the number of wage changes on a specific job to date, estimates of these constructs are of obvious interest. As Flinn and Heckman [1982] point out, the maximum likelihood estimator of the reservation wage given non-employment is simply the smallest wage offer observed to be acceptable. Because every acceptance decision censors wage observations drawn from an appropriately defined distribution, this observation generalizes as demonstrated below.

Consider a random sample of completed job spell durations and subsequent transitions to a new job. Suppose that the data include observations on $(\{w_{i-1},t_i\},k,\delta,w_0,t_0)$ for each element $h \in H$ of the sample. Equations (32)-(34) imply that the log likelihood for a sample of identical workers expressed as a function of reservation wage rates and offers is proportional to

$$(35)\qquad \ln L = \Sigma_{h\varepsilon H}(1-\delta^h)[\ln\lambda + \ln f_0(w_0^h)]$$

$$+ \Sigma_{h\varepsilon H}\delta^h[\ln\lambda - \lambda[1-F_0(\sigma_0)]t_0^h + \ln f_0(w_0^h)]$$

$$- \Sigma_n\Sigma_{h\varepsilon H_n}\sum_{i=1}^{n}\lambda[1-F_0(\rho_{i-1}(w_{i-1}^h))]t_i^h$$

$$+ \Sigma_n\Sigma_{h\varepsilon H_n}\delta^h \ln F_n(\sigma_n, w_{n-1}) + \Sigma_n\Sigma_{h\varepsilon H_n}\sum_{i=1}^{n-1}\ln f_i(w_i^h, w_{i-1}^h)$$

where the superscript denotes the observation association with a particular sample element and

$$(36)\qquad H_n = \{h \; \varepsilon \; H | k^h = n\}, \; n = 1,2,\cdots$$

is the subset of completed job spells that end after n wage subspells. The maximum likelihood estimates of the reservation wage rates and offers common to all observations maximize the function subject to the constraint that all observed wage rates are acceptable. Formal statements of these constraints follow:

(37.a) $\sigma_0 \leqslant w_0^h$ for all $h \; \varepsilon \; H$ such that $\delta^h = 1$.

(37.b) $\sigma_n \leqslant w_n^h$ for all $h \; \varepsilon \; H_n$ for all $n \geqslant 1$.

(37.c) $\rho_{n-1}(w) \leqslant w_0^h$ for all $h \; \varepsilon \; H_n$ and such that

$$\delta^h = 0 \text{ and } w_{n-1}^h = w \text{ and all } n \geqslant 1.$$

Notice that if at least one job spell ends with a transition to non-employment, i.e., $\delta^h=1$ for at least one element of H, then lnL is strictly increasing in σ_0 which implies that (37.a) is binding for some h. Hence, the maximum likelihood estimator is

$$(38.a) \qquad \hat{\sigma}_0 = \min_{h \in H} \{w_0^h \mid \delta^h = 1\} \text{ if the set is non-empty.}$$

Obviously, if no one in the sample ends a job spell with a transition to non-employment, the set defined on the right side of (38.a) is empty, then the sample provides no information about acceptable wage offers given non-employment and the reservation wage is not identified. Since lnL is also strictly increasing in each reservation wage given employment if and only if some worker quits to non-employment after wage spell n, the constraint (37.b) is binding for some h which implies that the maximum likelihood estimator of the reservation wage rates given employment state n is

$$(38.b) \qquad \hat{\sigma}_n = \min \{ w_n^h \} \text{ if the set is non-empty}$$

$$\text{and } \delta^h = 1 \text{ for some } h \in H_n, \ k = 1,2,\cdots.$$

The second condition as well as the first are needed for identification because, if no one makes a transition to non-employment from employment characterized by tenure state n, then the reservation wage estimate is any number less than the smallest observed accepted wage. Finally, because lnL is strictly increasing in every reservation offer provided that someone in the sample attained the specified wage-tenure state the constraint (37.c) is binding if a direct job to job transition is observed. Hence,

$$(38.c) \qquad \hat{\rho}_{n-1}(w) = \min\{w_0^h \mid h \in H_n,\ \delta^h = 0, \text{ and } w_{n-1}^h = w\}$$

if the set is not empty, $k = 1,2,\cdots.$

However, notice that if the set is empty say because no one in the sample leaves employment at or before tenure state k, i.e., H_i is empty for all $i < n$, then the maximum likelihood estimate of the reservation offer is the upper bound on the support of the distribution of initial offers. This implication is a consequence of the fact that none of the unobserved offers that are presumed to have arrived before tenure state n is attained were accepted.

The sample minimum has some clear advantages as an estimate of a reservation value. It is obviously consistent and is easy to compute. However, in any finite sample it is biased above the true value given homogeneity and is particularly sensitive to "unobserved heterogeneity". Given the size of data sets on job turnover available to the labor economist, small samples are not a problem were it not for unobserved differences across individual workers.

An alternative and more common approach for testing purposes would be to specify a functional form for the quit to an alternative job function and then estimate its parameters using observations on wage subspell durations contained in a sample of completed job histories. Specifically, given a sample of complete job histories, equation (32) implies a log likelihood of the sequences of wage subspell durations conditional on the wage sequences which is proportional to

$$\ln L_1 = \Sigma_n \Sigma_{h\varepsilon H_n} \sum_{i=1}^{n} [(1-\delta^h)\ln q_{n-1}(w^h_{n-1}) = q_{i-1}(w^h_{i-1})t^h_i] \tag{39}$$

given

$$q_i\ (w) = \lambda[1-F_0(\rho_i(w,b),m)]. \tag{40}$$

where m is the mean of the initial wage offer distribution and b is the non-employment benefit. One disadvantage of this approach is that any explicit parametric specification of the quit function is an implicit parametric specification of the initial wage offer distribution. Conversely, given a distribution function, the parametric form of the quit function is determined. For example, in the case of an exponential initial wage offer distribution,

$F_0(x,m) = \exp-\{x/m)$, equation (40) is $q_i(w) = \lambda\exp\{-\rho_i(w/m,b/m,1)\}$

by virtue of Corollary 2. Hence, a linear specification of the reservation offer function of the form

$$\rho_i(w/m,b/m,1) = \alpha_i + \beta w$$

implies

$$q_i(w) = \lambda\exp\{-\alpha_i-\beta w\} \tag{41}$$

assuming that different workers in the sample are identical with respect to (λ,m,b). Unique maximum likelihood estimates of the parameters $\{\alpha_i\}$ and β

can be calculated using equations (39). Furthermore, $\beta = 1/m$ given the general model, $\alpha_i = 0$ in the pure search on the job model, and $\alpha_0 = 0 > \alpha_i > \alpha_{i+1}$ in both the on the job training and job matching generalizations by virtue of Corollaries 3 and 4. Hence, a likelihood ratio test can be used to test for the absence of both of the latter two phenomena. If the test fails, one can also test the implication that it does so because α_i increases with i since estimates of the covariance matrix can be calculated from the Hessian of the likelihood function in the usual way. However, these tests are all conditional on the exponential specification of the initial wage offer distribution.

Of course, even if one were able to confirm the implication that reservation offers decrease with tenure and/or that reservation wage rates increase with tenure as measured by the number of wage changes to date by either of these methods, the problem of empirically distinguishing between the on the job training and the job matching hypotheses remains. Since the two hypotheses differ with respect to their implications about the distribution of job specific wage rates, basing a test directly on the observations of the wage rates earned on specific jobs suggests itself. Indeed, in the theoretical decision model analyzed in this paper, the on the job training hypothesis is characterized by the assumption that the sequence of wage rates received on a job is a sub-martingale with diminishing drift while the job matching hypothesis is characterized by the assumption that the same sequence is a martingale with diminishing variance. In principle, general methods exist that permit testing of these alternatives. However, the theoretical decision model implies that one must take appropriate account of the fact that on the job wage observations are censored, i.e., only those that exceed the employment state contingent reservation wage are observed.

Since low wage observations are censored by the decision to quit to non-employment, observed rates of on the job wage growth over estimate the effect of specific job experience on productivity. Furthermore, if reservation wage rates increase with tenure as predicted, then the extent of the bias increases with tenure. Since these biases might otherwise lead one to reject both the on the job training and the job matching hypotheses when one or the other or some combination of the two are true, an appropriate correction for censoring is necessary. Equation (32) implies that the maximum likelihood estimates of the parameters of the distributions of the sequence of on the job wage rates given the general model are corrected for this form of censoring. Specifically, given a sample of completed job histories, the log likelihood of the observed sequences of wage rates conditional on the wage subspell durations is proprotional to

$$(42) \qquad \ln L_2 = \Sigma_n \Sigma_{h\varepsilon H_n} \left[\sum_{i=1}^{n-1} \ln f_i(w_i, w_{i-1}^h) + \delta^h \ln F_n(\sigma_n, w_{n-1}^h) \right].$$

The information that each worker who quits to non-employment sees an unacceptable wage at the end of the job spell which is not observed by the econometrician is contained in the second term between the square brackets. It is of interest to illustrate its effect on the estimates of the on the job rates of productivity growth with an example.

For the sake of the example, assume an exponential distribution of the next wage given the previous one so that

$$(43) \qquad F_i(x,w) = \exp\{-x/\beta_i w\}.$$

Note that β_i is the mean of the unconditional distribution of the ratio of the next wage to the current wage given tenure i; i.e., it equals one plus the rate of expected growth in productivity experienced as a consequence of a transition from tenure i to tenure i+1. Hence, the on the job training hypothesis would imply that β_i exceeds unity but declines with i were the specification correct. Substituting from (43) into (44), one obtains

$$(44) \qquad \ln L_2 = \Sigma_n \Sigma_{h\varepsilon H_n} \Big[\sum_{i=1}^{n-1} \{\ln(w_i^h/\beta_i w_{i-1}^h) - w_i^h/\beta_i w_{i-1}^h\}$$

$$+ \delta^h \ln[1-\exp\{\sigma_n/\beta_n w_{n-1}^h\}\}]$$

The maximum likelihood estimate β_n is a solution to the following first order condition:

$$\Sigma_{h\varepsilon H(n)}[\beta_n - w_n^h/w_{n-1}^h] + \Sigma_{h\varepsilon H_n} \delta^h[\sigma_n/w_{n-1}^h]/[1-\exp\{\sigma_n/\beta_n w_{n-1}^h\}] = 0$$

where $H(n) = \{h\varepsilon H \mid k^h \geqslant n\}$ is the set of workers who attain tenure n before ending their job spell. Because the last term is positive if any one in the sample quits to non-employment at tenure n, the estimate of β_n is less than one plus the sample rate of growth in observed wage rates as asserted earlier. Because the last term is also a strictly increasing function of the reservation wage σ_n and decreasing function of β_n, the extent to which the sample rate of growth in observed wage rates exceeds productivity growth increases with tenure if the latter decreases with tenure as the on the job training hypothesis implies.

Appendix A: The Existence and Uniqueness of the Value Function

All the turnover models studied in this paper can be regarded as a special case of the following general structure. At any point in time the worker's employment "state" is characterized by a wage-tenure pair $(w,n) \in WxN$ where W is a subset of the positive reals and N is the set of non-negative integers. The worker receives outside offers at the given Poisson rate $\lambda \geq 0$, which is independent of the worker's state. Each initial wage offer is a random draw from the stationary wage offer c.d.f. $F_0(\cdot,m):W \rightarrow [0,1]$. New job specific wage rates arrive via a Poisson c.d.f. $F_n(\cdot,w):w \rightarrow [0,1]$ which generally depends on the worker's current wage and tenure, the pair (w,n), as the notations indicates. Finally, the worker lives forever, discounts future income streams at the rate $r > 0$, and moves from one job to another when the opportunity arise so as to maximize expected future discounted income flows, wealth.

Given this structure, we are seeking conditions that guarantee the existence of a value of employment function, a map $v(\cdot,u,m):WxN \rightarrow R$ from employment states to wealth given any finite value of non-employment and mean initial wage offer pair (u,m). As established in the text, equation (14), the value function is a fixed point of a transformation which maps the set of functions defined on WxN into itself.

Burdett's [1978] search on the job model can be interpreted as the special case in which subsequent wage rates equal the current wage with probability one or, equivalently, the case in which no new job specific wage rates arrive while on a given job. The on the job training model of the type popularized by Mincer (1974) is the case in which specific wage rates are expected to increase with tenure but at a diminishing rate in the sense that the c.d.f. $F_n(x,w)$ increases in n. Finally, in the case of Jovanovic's

[1979,1984] matching model, new job specific wage rates are interpreted as adjustments made in response to new information gathered about the worker's productivity. In its most general form, the hypothesis implies that the mean of $F_n(x,w)$ is w for all n but that the "mean preserving spread" decreases with n as a consequence of the fact that the current estimate of the worker's true job specific, the wage, becomes more precise as more observations on a worker's realized productivity are accumulated.

The general theory of turnover studied in this paper has meaning if and only if the transformation T defined by (14) has a unique fixed point under conditions that are consistent with these various interpretations. the purpose of this section is to verify that such is the case.

The usual approach used to prove existence and uniqueness of the value function in stochastic dynamic programming is due to Blackwell [1965]. A direct application of the approach requires that one establish that the functional equation derived using Bellman's principle, equations (12) in our case, defines the value function as a fixed point of contraction map, T defined by (14) in our case, from the space of bounded functions into itself. Although the right side of (14) does define a contraction, it does not generally map bounded functions into the set of bounded functions simply because w/r is not bounded on the reals. Of course, one could apply the approach directly were one willing to make the realistic assumption that the set of possible wage rates, W, is bounded. However, there are two problems with taking this approach. First, the assumption implies that the support of all the conditional distributions of all future wage rates are bounded. In empirical work, it will generally be convenient to approximate these distributions by standard parametric forms, e.g., log normal or Pareto, which do not have this property. Doing so is appropriate and useful only if an associated value function exists. Second, because the model supposes an infinite time horizon and because the support of the distribution of subsequent wage rates on a specific job depends on the current wage, the set of possible wage rate at any date in the indefinite future need not be bounded even under the assumption that the support of each in the infinite sequence of job specific wage distributions is bounded.

The proof presented in this section of the paper uses a modifications of Blackwell's approach. The "trick" is to define a new function

$$\text{(A.1)} \qquad f_n(w) = v_n(w) - k_n w/r \quad \text{for every } (w,n) \; \varepsilon \; W\text{x}N$$

where $\{k_n\}$ is a sequence of constants to be determined. Given any sequence, a function $v{:}W\text{x}N \to R$ solves (12) if and only if a function $f{:}W\text{x}N \to R$ solves

$$f_n(w) + k_n w/r = (1-\beta_1-\beta_2)w/r$$

$$+ \beta_1 \int_W \max[f_0(x)+k_0 x/r, f_n(w)+k_n w/r] dF_0(x,m)$$

$$+ \beta_2 \int_W [f_{n+1}(x)+k_{n+1}x/r, u] dF_{n+1}(x,w).$$

In other words, f is a fixed point of the transformation M defined by

$$Mf_n(w) = (1-\beta_1-\beta_2)w/r - k_n w/r + \beta_1 k_n w/r + \beta_2 k_{n+1} w/r$$

$$+ \beta_1 \int_W \max[f_0(x)+k_0 x/r - k_n w/r, f_n(w)] dF_0(x,m)$$

$$+ \beta_2 \int_W \max[f_{n+1}(x)+(x-w)k_{n+1}/r, u-k_{n+1}w/r] dF_{n+1}(x,w).$$

Since max[a+c,b] = c + max[a,b−c], this equation can be rewritten as

(A.2) $\quad Mf_n(w) =$

$$[1 - \beta_1 - \beta_2 - k_n + \beta_1 k_n + \beta_2 k_{n+1}(1+g_{n+1}(w))]w/r$$

$$+ \beta_1 \int_W \max[f_0(x)+k_0 x/r - k_n w/r, f_n(w)]dF_0(x,m)$$

$$+ \beta_2 \int_W \max[f_{n+1}(x), u - k_{n+1}x/r]dF_{n+1}(x,w)$$

where

(A.3a) $\quad \beta_1 = \lambda/(r+\lambda+\eta)$

(A.3b) $\quad \beta_2 = \eta/(r+\lambda+\eta)$

by virtue of equation (13) and

(A.4) $\quad q_n(w) \equiv \int_W \frac{x-w}{w} \; dF_n(x,w), \quad n = 1,2,\cdots$

is the expected proportional change in the job specific wage at arrival n. Consequently, if one can find a sequence $\{k_n\}$ such that M has a unique fixed point, then the function v constructed using the sequence, fixed point, and (A.1) is the only solution to (12).

Theorem 2: Given $F_0(\cdot,m):W \to [0,1]$ and $F_n(\cdot,w):W \to [0,1]$ for all $(w,n) \in W\mathrm{x}N$, a sequence of constants $\{k_n\}$ can be found such that $f = Mf$ exists and is unique if for all $(w,n) \in W\mathrm{x}N$

(A.5a) $F_n(x,w)$ is continuous and decreasing in w and

$g_n(w)$ is continuous and non-increasing in w,

(A.5b) $0 \leq g_n(w) < \infty$

and a finite number n^* exists such that

(A.5c) $g_n(w) < r/n$ for all $n \geq n^*$.

Proof: Because (A.2) implies that M is increasing in the sense that $f_n(w) \geq h_n(w)$ implies $Mf_n(w) \geq Mh_n(w)$ and because $||M[f+c]|| \leq ||Mf|| + (\beta_1+\beta_2)c$, where $0 \leq \beta_1+\beta_2 < 1$ from (A.3), for every constant c where $||\cdot||$ represents the sup norm, the map M satisfies Blackwell's [1965] sufficient conditions for a contraction. Hence, it suffices to show that a sequence of constants $\{k_n\}$ exists such that M given the sequence maps the set a Banach space into self.

The remainder of the argument is presented in two parts. First, we show that any sequence satisfying the following conditions implies that $Mf_n(w)$ is continuous, decreasing in w, and bounded above on W, a subset of the positive reals, if $f_n(w)$ has the same property by virtue of (A.5a):

(A.6a) $\infty > k_n \geq 0$

(A.6b) $1-\beta_1-\beta_2- k_n(1-\beta_1) + \beta_2(1+g_{n+1}(w))k_{n+1} \leq 0.$

The set of all such functions is a Banach space. Second, we show that (A.5b) and (A.5c) implies that such a sequence exists.

That $Mf_n(w)$ is continuous and decreasing given (A.6) and (A.5a) if $f_n(w)$ is continous and decreasing is an obvious implication (A.2). That $Mf_n(w)$ is bounded above on W given that $f_n(w)$ is bounded above follows as a consequence of the following implication of (A.2), (A.5a), (A.6) and the fact that W is a subset of the non-negative reals:

$$Mf_n(w) \leq \beta_1 \int_W f_0(x) dF_0(x,m) + \beta_1 k_0 m/r + \beta_1 f_n(w)$$

$$+ \beta_2 \int_W f_{n+1}(x) dF_{n+1}(x,w) + \beta_2 u.$$

Specifically, the first term on the right side of (A.2) is decreasing in w on W given (A.5a) and (A.6b) and hence takes its maximum value at $w = 0$. The remainder of the inequality is implied by (A.6a), that wage rates are non-negative, and the fact that $\max[y,z] \leq y + z$.

Finally, the sequence generated by the following difference equation satisfies the conditions (A.6) by virtue of (A.3), (A.5b) and (A.5c).

$$\text{(A.7a)} \qquad k = k_n = (1-\beta_1-\beta_2)\ (1-\gamma)^{-1} \quad \text{for all } n \geq n^*$$

$$\text{(A.7b)} \qquad k_{n-1} = (1-\beta_1-\beta_2)\ (1-\beta_1)^{-1} + \beta_2(1+g_n)\ (1-\beta_1)^{-1} k_n \text{ otherwise.}$$

where

$$\text{(A.8a)} \qquad 0 \leq \gamma \equiv \beta_1 + \beta_2(1 + \sup_{w,n\geq n^*} [g_n(w)]] < \frac{\lambda + \eta(1+r/\eta)}{r + \lambda + \eta} = 1$$

and

(A.8b) $\quad \infty > g_n \equiv \sup_{w} [g_n(w)] \geqslant 0.$

<u>Corollary</u>: A unique value function $v(\cdot,u,m):WxN \to R$ exists for every finite (u,m).

Since the average length of the spells between the arrivals of wage increments is $1/\eta$, condition (A.5c) requires that the expected rate of exogenous on the job exponential wage growth per unit period fall below the interest rate eventually as the worker accumulates tenure. This requirement is a standard one in capital theory for the existence of finite discounted future income streams. However, establishing that it is sufficient for the existence is a new result in the ltierature on stochastic turnover theory.

Appendix B: Bayesian Models of Learning about the Job

The analysis of the job matching hypothesis in the text is based on three assumptions. First, the mean of the worker's productivity at the next arrival given the sequence of observations to date is equal to the current wage; i.e., the stochastic wage generation process is a Martingale. Second, the conditional distribution of the next wage on a job is stochastically increasing in the current wage. Third, the current wage as a conditional estimate of the next wage becomes more precise in the sense that the "mean preserving spread" of the sequence of wage distributions decreases with n given the mean w. Formally,

(B.1) $\quad \int_W x dF_n(x,w)dx = w$

(B.2) $\quad F_n(x,w) \leqslant F_n(x,w)$ for all $x \varepsilon W$ if $w \geqslant w$

and

(B.3) $\int_0^y F_n(x,w)dx \;>\; \int_0^y F_{n+1}(x,w)dx$ for all $y \in W$.

Both the plausibility and meaning of this structure is best defended by example:

Example 1: Learning about the mean of a normal distribution: Jovanovic [1979,1984] assumes that the sequence of productivity observations $\{x_i\}$ are i.i.d. normal with common unknown mean μ, the worker's true job specific productivity, and known variance σ^2. the prior distribution of μ is normal with known mean m and known variance s^2. The standard sequential Bayesian estimation problem is to form an estimate of μ using the sequence of observations to date and then to update the estimate as more observations are made. Under the normality assumptions, it is well known (see DeGroot [1972, p.167]) that the posterior distribution of μ given the observed sequence of past observations is normal with mean

$$(\text{B4.a})\qquad w_n \equiv E\{\mu|x_1,\cdots,x_n\} = [m/s^2 + (\sum^n x_i)/\sigma^2]v_n^2$$

and variance

$$(\text{B4.b})\qquad v_n^2 = E\{(\mu - w_n^2) \mid x_1,\cdots,x_n\} = 1/[1/s^2 + n/\sigma^2].$$

In the turnover model, the wage paid at any date is assumed equal to the current estimate of the worker's true productivity as the notation in (B4) suggests. To decide whether to switch jobs were an alternative available, the worker needs an estimate of the future time path on the current job. Under the assumption of rational expectations, that estimate is based on the wage generation rule (B4.a). For this purpose, it is permitted and convenient to use the current wage as the sufficient statistic along with the number of productivity observations rather than the accumulated productivity. Because the equations of (B4) imply

$$\text{(B5)} \qquad w_n = w_{n-1} + [1-v_n^2/v_{n-1}^2][x_n - w_{n-1}],$$

and because x_n is normal with mean w_{n-1} and variance $\sigma^2+v_{n-1}^2$, the conditional distribution of the wage paid at the next arrival of a productivity observation given that the current wage is normal with mean

$$\text{(B6.a)} \qquad E\{w_n \mid w_{n-1} = w\} = w,$$

and variance

$$\text{(B6.b)} \qquad s_n^2 \equiv E\{[w_n-w]^2 \mid w_{n-1} = w\}$$

$$= \{1-v_n^2/v_{n-1}^2]^2[v_{n-1}^2 + \sigma^2] = v_n^2 \, v_{n-1}^2/\sigma^2,$$

by virtue of (b4.b). In sum, the conditional distribution of the next wage given the current wage obviously satisfies (B1) by virtue of (B6.a). In the case of the normal family, a higher mean given the variances implies stochastic dominance, condition (B2). Further, holding the mean fixed, the variance is a measure of mean preserving spread in the normal distribution case. Hence, condition (B3) is satisfied by virtue of and (B6.b).

In general, the assumption that the current wage is an unbiased estimate of productivity given current information can be viewed as equivalent to the assumption that expectations are rational. The further restriction that the mean is also an indicator of stochastic dominance holds in many cases but need not be general. Finally, in the example, dispersion in the distribution of future wage rates falls with the number of observations as a consequence of the fact that the variance in the estimate of the worker's true productivity falls with the sample size. The general rationale is that the uncertainty about the worker's true productivity can be expected to decrease with the amount of information as measured by the sample size. The following example provides another case for which all three restrictions hold.

Example 2. Learning about the probability of success: Suppose that the observations on "productivity" are simply "success" or "failure" in performing the job task so that $\{x_i\}$ is a sequence Bernoulli trials with unknown probability of success π. Assuming a Beta prior for π with parameters α and β, where the mean is $m = \alpha/(\alpha+\beta)$, the posterior distribution given n observation is Beta with mean

$$(B7)\qquad w_n = [\alpha + \sum_1^n x_i]/[\alpha+\beta+n] = w_{n-1} + [x_n - w_{n-1}]/[\alpha+\beta+n].$$

(Again, see DeGroot [1972, p.160].) The distribution of x_n given w_{n-1} is Bernoulli with probability of success w_{n-1}. Since the random variable w_n given $w_{n-1} = w$ has mean w, it takes on only one of two values, the larger with probability w, and the difference between the two values diminishes with n, the assumptions of (B1)-(B3) are clearly satisfied.

Notice that in both examples, the distribution of the initial wage offer on any job is the prior mean with probability one. This inference follows by virtue of the assumption that no new information is obtained until the arrival of the first productivity observation. However, one can easily generalize the model by supposing that each prospective new worker is given a test and that this information is incorporate along with prior information about a worker productivity based on observable characteristics, such as education, in the initial wage offer. Under the reasonable assumption that the expectation of the test score is equal to the prior estimate of the worker's productivity, the distribution of the initial wage has mean m.

FOOTNOTES

[1]This argument requires that the map M have a unique fixed point. The theorem in Appendix A establishes that M is a contraction which implies uniqueness a unique fixed point if the set of functions at issue is a complete metric space. The set of increasing continuous functions is such a space. The set increasing continuous functions is such a space.

[2]The statement is valid provided that the set of functions is a complete metric space in each case. See footnote 1.

[3]In other words, the wage offer distributions for two different workers differ only by a scale factor. This assumption implies that the variance of the log wage is independent of the mean as in the log normal case.

REFERENCES

Bartel, A.P., and G.J. Borjas [1981]: "Wage Growth and Job Turnover: An Empirical Analysis," in S. Rosen, ed., Studies in Labor Markets, University of Chicago Press, Chicago.

Bellman, R., [1957]: Dynamic Programming. Princeton University Press, Princeton, N.J.

Blackwell, D., [1965]: "Discounted Dynamic Programming," Annals of Mathematical Statistics, 36, 226-235.

Burdett, K., [1978]: "Employee Search and Quits," American Economic Review, 68, 212-220.

Flinn, C.J., [1983]: "Wages and Job Mobility of Young Workers," IRP Discussion Paper #728-83, University of Wisconsin-Madison.

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

Gronau, R., [1974]: "Wage Comparison--A Selectivity Bias," Journal of Political Economy, 82, 1119-1144.

Heckman, J.J. and B. Singer [1982]: "The Identification Problem in Econometric Models for Duration Data" in W. Hildenbrand, ed., Advanced in Econometrics, Cambridge, U.K., Cambridge University Press, 39-77.

Johnson, W., [1978]: "A Theory of Job Shopping," Quarterly Journal of Economics, 92, 261-277.

Jovanovic, B., [1979a]: "Job-Matching and the Theory of Turnover," Journal of Political Economy, 87, 972-90.

Jovanovic, B., [1979b]: "Firm-Specific Capital and Turnover," Journal of Political Economy, 87, 1246-1260.

Jovanovic, B., [1984]: "Matching, Turnover and Unemployment," Journal of Political Economy 92, 108-122.

Kalbfleish, J., and R. Prentice [1980]: The Statistical Analysis of Failure Time Data, Wiley, New York.

Kiefer, N., and G. Neumann [1979a]: "An Empirical Job Search Model with a Test of the Constant Reservation Wage Hypothesis," Journal of Political Economy, 87, 69-82.

Kiefer, N., and G. Neumann [1979b]: "Estimation of Wage Offer Distributions and Reservation Wages" in Lippman and McCall [1979].

Kiefer, N., and G. Neumann [1981]: "Individual Effects in a Nonlinear Model: Explicit Treatment of Heterogeneity in the Empirical Job Search Model," Econometrica, 49, 1981.

Marshall, R.C., [1983]: "The Effects of Job Specific Human Capital Accumulation on Wages: A Censored Regression Approach", unpublished paper, University of Virginia.

Mattila, J.P. [1974]: "Job Quitting and Frictional Unemployment," American Economic Review 64, 235-240.

Mincer, J., [1974]: Schooling, Experience, and Earnings, Columbia University Press, New York.

Mincer, J., and B. Jovanovic [1981]:"Labor Mobility and Wages" in S. Rosen, ed., Studies in Labor Markets, NBER Conference Report, University of Chicago Press, Chicago.

Mortensen, D.T., [1978]: "Specific Capital and Labor Turnover," Bell Journal of Economics, 9, 572-586.

Parsons, D.O., [1977]: "Models of Labor Turnover, A Theoretical and Empirical Survey," in R. Ehrenberg, ed., Research in Labor Economics, vol. 1, 185-225 JAI Press, Greenwich, Conn.

Rothschild, M., [1973]: "A Two-Armed Bandit Theory of Pricing," Journal of Economic Theory, 9, 185-202.

Rothschild, M., and J. Stiglitz, [1970]. "Increasing Risk I: A Definition," Journal of Economic Theory, 2, 225-243.

Viscusi, K., [1979]: "Job Hazards and Worker Quit Rates: An Analysis of Adaptive Worker Behavior," International Economic Review, 20, 29-58.