FACTORS AFFECTING THE OUTPUT
AND QUIT PROPENSITIES
OF PRODUCTION WORKERS

Roger Klein

Richard Spady

Andrew Wiess

Factors Affecting the Output and Quit
Propensities of Production Workers

ABSTRACT

We have used a proprietary data set of newly hired semi-skilled production workers at one location of a large unionized firm to investigate several issues in labor economics. This data set is unique in several respects: the workers in our sample faced the same wage schedules, had the same promotional opportunities, the same job tenure (zero), similar working conditions, and had jobs for which we were able to record their physical output. We analyze these data by formulating a simultaneous equation model to explain wages, output, education, and a worker's quit decision. The model is estimated by maximum likelihood and subjected to a variety of specification tests. Such tests include a comparison of the standard error estimates that form the basis for White's information test, and White's version of a Hausman specification test.

Our principal findings are:

1. Individuals that choose more education than we would expect from their observed characteristics have lower than expected quit propensities. We argue that this low quit propensity is one of the unmeasured (and unobserved) attributes that sorting models posit are correlated with education and hence distort the usual estimates of rates of return to education.

2. The performance of non-whites in our sample was no lower than that of whites. However, on their previous jobs non-whites received lower wages than did whites.

3. The output per hour of males in our sample was higher than that of females; however, we were unable to conclude from our data whether these productivity differences could explain the higher wages received by men on their previous jobs. Moreover, this output difference may be transitory and may diminish with on-the-job learning.

4. The expected value of alternative wages had a positive (but not statistically very significant) effect on quit rates. Workers with better alternative opportunities were more likely to quit (all workers had the same opportunities on their current job).

5. Finally we found that workers with high output levels were more likely to quit than were workers with average output levels.

Roger Klein
Bell Communications
  Research
435 South Street
Morristown, NJ 07960

Richard Spady
Bell Communications
  Research
435 South Street
Morristown, NJ 07960

Andrew Weiss
Bell Communications
  Research
435 South Street
Morristown, NJ 07960

FACTORS AFFECTING THE OUTPUT AND QUIT PROPENSITIES OF PRODUCTION WORKERS

Roger Klein, Richard Spady, Andrew Weiss

Bell Communications Research
435 South Street
Morristown, New Jersey 07960

1. INTRODUCTION

We have used a proprietary data set of newly hired semi-skilled production workers at one location of a large unionized firm to investigate several issues in labor economics. This data set is unique in several respects: the workers in our sample faced the same wage schedules, had the same promotional opportunities, the same job tenure (zero), similar working conditions, and had jobs for which we were able to record their physical output. We analyzed these data by formulating a simultaneous equation model to explain wages, output, education, and a worker's quit decision. The model was estimated by maximum likelihood and subjected to a variety of specification tests. The tests included a comparison of the standard error estimates that form the basis for White's information test, and White's version of a Hausman specification test.

Our principal findings are:

1. Individuals that choose more education than we would expect from their observed characteristics have lower than expected quit propensities. We argue that this low quit propensity is one of the unmeasured (and unobserved) attributes that sorting models posit as being correlated with education and hence distort the usual estimates of rates of return to education.[1]

2. The performance of non-whites in our sample was no lower than that of whites. However, on their previous jobs non-whites received lower wages than did whites.

3. The output per hour of males in our sample was higher than that of females; however, we

were unable to conclude from our data whether these productivity differences could explain the higher wages received by men on their previous jobs. Moreover, this output difference may be transitory and may diminish with on-the-job learning.

4. The expected value of alternative wages had a positive (but not statistically very significant) effect on quit rates. Workers with better alternative opportunities were more likely to quit (all workers had the same opportunities on their current job).

5. Finally we found that workers with high output levels were more likely to quit than were workers with average output levels. We argue that the high performance workers were likely to have quit because of the lack of any individual rewards for high output — neither pay nor promotional opportunities were significantly affected by performance. Those workers presumably sought jobs with firms that rewarded individual performance.

A firm specific data set of the sort we assembled has several attractive features. First, all of the workers in our sample had bench-assembly jobs; consequently, we were able to obtain accurate measures of their physical output. Second, since these workers were employed in the same plant, the normalizations performed by the industrial engineers to determine equivalent quantities of output for workers on different jobs were likely to be fairly accurate.

Third, because the promotional opportunities, working conditions, and wage schedules of the workers were identical, we did not have to separate firm specific and job specific effects from person specific effects. For example, in cross-section analyses of individuals on different jobs, it has been found that better educated workers have lower quit rates. Without separating individual from firm and job specific effects, it would be impossible to know whether to attribute this finding to a lower quit propensity of better educated workers or to the nature of the jobs performed by better educated workers.

Fourth, by only considering newly hired workers we avoid some of the simultaneous equation biases present when tenure is included as an "independent" variable in a quit equation.

Finally, by obtaining information about workers from the personnel files of the firm rather than from interviews with the workers, we are likely to have less measurement error in the usual demographic characteristics such as education, age, experience, and tenure than is customary in survey data.[2] For example, we were able to compute precisely how long workers were employed before they quit: survey data on questions of that sort are notoriously inaccurate.

On the other hand our data have two major shortcomings, the first of which is sample selection bias. For example, semi-skilled production workers with post-secondary schooling are not likely to be typical of individuals with those education levels (though fewer than ¼ of our sample consists of such individuals).

As is argued in the body of the paper, sample selection bias is not a serious problem in our data. Moreover, since we did not have data on non-applicants, corrections of this bias would be extremely sensitive to the unverifiable distributional assumptions that we would be forced to make. Consequently, rather than attempting to correct for sample selection bias, we focused our attention on explaining empirical regularities within our sample. (However, many of the relationships we sought to explain within our data, such as the higher wages of better educated workers, males, and whites, are also found in larger, standard data sets.)

Although our methodology is not completely applicable to other data sets, we argue that it is not unreasonable to cautiously generalize our results as providing partial explanations for the empirical regularities found in standard data sets. The sample selection biases that may hinder these generalizations are discussed at greater length in section 3.5.

The second drawback to our data pertains to our observations on wages. Because all the workers were receiving the same wage on their current job, we focused on explaining wage differences on the previous jobs held by these workers. Of course these are jobs from which the workers either quit or were laid-off. Thus it is difficult to know how to interpret those wages. For workers who were laid-off from their previous job, specification diagnostics indicated that

the model was misspecified.[3] Those workers were omitted from the analysis. For the sub-sample of workers who quit their previous job, the model performed well under a variety of specification tests (section 5). Accordingly, we restricted the sample to individuals who quit their previous job, and assumed that the wage for such workers on their previous job was a reasonable estimate of the market wage for individuals with their observed characteristics.

The format of the paper is as follows: section 2 describes the data, section 3 presents the model that we estimated, section 4 presents the estimation procedure, and section 5 discusses in detail the model specification tests we performed. In addition to a variety of likelihood ratio tests, such tests included White's information test, White's version of a Hausman specification test, and a test of the assumed error distribution in the discrete quit education. The latter test is important, because the estimated quit equation is much more sensitive to its assumed error distribution than is the case for the remainder of the model. These tests validated our preferred specification. For this preferred specification, in section 6 we describe the results and implications of the model. In section 7 we discuss the limitations to this study and directions of future research.

2. THE DATA

The data, which are summarized in Table 1, consist in part of detailed biographical information for 463 semi-skilled production workers hired at a single manufacturing facility in 1977 and 1979.[4] All of these workers were employed at the time they were hired. We obtained records of the output per hour of these workers and whether or not they quit during their first six months on the job. The biographical information which we obtained from personnel files included the workers' sex, age, race, years of education, marital status, starting and ending dates on their four previous jobs and their pay on those jobs, and their score on each half of the dexterity test used in screening applicants (the first half involved inserting small pins in sleeves with tweezers, while the second half involved screwing screws into sleeves). We also obtained an

estimate by the industrial engineering staff of the time required to learn each job to which the workers were assigned. In estimating the model below, job complexity was defined as the logarithm of this predicted training requirement.

Plant supervisors kept detailed records of the physical output per hour of newly-hired workers. The industrial engineers took considerable care in scaling output across jobs to ensure that the expected "measured output" of a worker who was experienced on any pair of jobs would be the same on the two jobs. These scaling factors are frequently audited by the corporate headquarters of the firm which finds only trivial discrepancies. In addition, the industrial engineers compute learning curves for each job to measure the number of weeks the average worker takes to learn the job and the percentage of the expected output for an experienced worker that a newly-hired worker would be expected to achieve during his first month on the job. This percentage is referred to as expected output. (The learning curves are not audited and our data suggest that the industrial engineers over-estimate the difficulty of learning complex jobs. We adjusted for those over-estimates.) Measured output during the first month was divided by expected output to obtain a measure of normalized output per hour.

During their first month on the job, each worker was paid on the same convex, piece-rate schedule, where the piece-rate was computed from their normalized output. Once a worker achieved 83.5% of the expected output for an experienced worker on the job, the worker was assigned to a pay group. All members of a pay group received the same pay, which was based on the output of the group. The average group had 126 members. Therefore, there was little monetary incentive for individuals on group incentives to achieve high output levels: group incentives were roughly equivalent to time based pay. Because of this lack of incentives, or because of social pressures by workers and managers, there was very little difference in the output of experienced workers. Consequently, we only used the workers' normalized outputs during their first month on the job in the empirical work described below.

Finally, we were fortunate that workers were randomly assigned to jobs, received the same pay schedule, and had the same probability of being promoted (nil). Although each worker had to achieve a passing grade on the dexterity tests before being hired, none of the observed characteristics of workers were used in assigning workers to jobs. We checked on whether this corporate policy of random job assignment was followed in practice by regressing job complexity against observed demographic characteristics. None of the estimated coefficients were economically or statistically significant. The direct correlations between various observed characteristics and job complexity are included in Table 1. As can be seen, education, experience, previous wage and score on the screws section of the dexterity test are all negatively and weakly correlated with the training required to learn ones assigned job. This is consistent with the assertion by the personnel office that job assignments were random. To explain the random assignment, the head of the personnel office at the plant told one of the authors that "EEOC (Equal Employment Opportunity Commission) rules dictate that any worker who satisfies the hiring standards can perform any job."

3. THE MODEL

The endogenous variables in our system of equations are: pay on the individual's previous job (as we pointed out, lifetime pay on the current job varies little across workers), years of schooling, normalized output per hour on the current job, and a zero-one variable indicating whether the individual quit during his first six months on the job. Clearly, some of these endogenous variables affect others. For example, education is likely to affect previous pay, and may affect quit propensity and output per hour. We also allowed for correlations among the error terms in the equations we estimated. (As we discuss below, the economic implications of these correlations are interesting.) Therefore, we estimated the equations in the model simultaneously by the method of maximum likelihood.

3.1 Quit Propensity

The model that we estimated can best be described by beginning with the quit equation. There is a large literature in psychology journals studying why workers quit their jobs (see Mobley *et al*, Knapp *et al* or Arnold and Feldman). As expected, the main reasons offered are dissatisfaction with the current job, superior opportunities elsewhere, and individual specific preferences for moving.

To model these effects for worker i on job k at time t, denote $\widetilde{W}_{ik}^R(t)$ as his real income, inclusive of fringe benefits, and $N_{ik}(t)$ as a vector of idiosyncratic benefits. Let utility be additively separable in $\widetilde{W}^R$ and N and exhibit diminishing marginal utility of income. Specifically, we assume that utility for worker i in job k at time t, $V_{ik}(t)$, is given as:

$$V_{ik}(t) = \beta_1 Ln \widetilde{W}_{ik}^R(t) + \beta_N N_{ik}(t) . \tag{1}$$

Next, denote M_i as i's costs (real and psychic) of a job change (housework and leisure are counted as jobs). Then, worker i quits if the discounted utility differential between the best alternative job (A) and the current situation (C) exceed M_i. Accordingly, in a world with perfect foresight, with a_i as worker i's age, δ the discount rate, and $\Delta N_i \equiv N_{iA} - N_{iC}$ (the difference in idiosyncratic benefits), for worker i to quit $(Q = 1)$, we have:

$$Q = 1 \text{ iff } \int_0^{65-a_i} e^{-\delta t} \left\{ \beta_1 \left[Ln \overline{W}_{iA}^R(t) - Ln \overline{W}_{iC}^R(t) \right] + \beta_N \Delta N_i(t) \right\} dt > M_i . \tag{2}$$

We are assuming that all workers retire at 65. At the time our data were collected, 65 was the mandatory retirement age for the workers in our sample. The modal retirement age in the U.S. is currently around 62. Because our sample is young and its average age is 23, we are discounting the future and our results are not sensitive to small changes in the assumed retirement age.

Since neither the authors nor the workers in the sample have perfect foresight concerning future wages or fringe benefits, in order to estimate (2) it was necessary to assume that for each

job workers would expect the ratio of fringe benefits to wages to remain constant over time. This ratio varies across jobs and is denoted by α_A for the alternative job and α_C for the current job. We also assumed that workers expect real wages to grow at the same rate $g(t)$ in all jobs, and that the difference in idiosyncratic benefits is constant. Letting W_{ik}^R denotes real base period wages exclusive of fringe benefits for worker i in job k, from the above assumptions:

$$\overline{W}_{iA}^R(t) = g(t) W_{iA}^R \quad (1 + \alpha_A)$$

$$\overline{W}_{iC}^R(t) = g(t) W_{iC}^R \quad (1 + \alpha_C) \tag{3}$$

$$\Delta N_i(t) = \Delta N_i \ .$$

Substituting (3) into (2) and letting $f(a_i, \delta) \equiv \dfrac{1 - e^{-\delta(65 - a_i)}}{\delta}$, we have:

$$Q_i = 1 \text{ iff } f(a_i, \delta) \left[\beta_1 [\beta_f + \text{Ln}(W_{iA}^R / W_{iC}^R)] + \beta_N \, \Delta N_i \right] > M_i \ , \tag{4}$$

where $\beta_f = \text{Ln} \left(\dfrac{1 + \alpha_A}{1 + \alpha_C} \right)$.

It now remains to explain how the variables in (4) are measured. We did not have a direct measure of the best alternative wage offered the worker. We did, however, know the nominal wage the worker received on his previous job, which we denote by W_{ip}^N. By adjusting W_{ip}^N, we developed a measure of the worker's best alternative real wage. First, we adjusted W_{ip}^N for local demand conditions and price effects (at the worker's job location where his previous wage was received). To perform this adjustment, we divided W_{ip}^N by the average manufacturing wage at the time and place the worker was previously employed. W_i denotes the resulting real wage variable.

As a further adjustment towards measuring real alternative wages, we adjusted W_i for tenure and for age and experience. These latter variables may have affected previous wages by enabling the worker to have searched longer over the distribution of job offers. To perform this adjustment, we estimated a wage equation (see (3.4)) of the form:

$$\operatorname{Ln} W_i = X_{1W}\, \beta_{1W} + X_{2W}\, \beta_{2W} + u_w \tag{5}$$

where X_{2W} consists of tenure, age, and experience, X_{1W} consists of other factors defined below, and u_w is the error term. We estimated this wage equation at the same time as the quit equation. We used (5) to define a previous wage, adjusted for tenure, age, and experience:

$$\operatorname{Ln} W_i^* = \operatorname{Ln} W_i - X_{2W}\, \beta_{2W} \ . \tag{6}$$

While we might now presume that W_i^* represents alternative real wages for a typical worker, some further adjustment is required for very high or very low productivity workers. Let B be a breakpoint (estimated below at 42% above the expected output level) that partitions high from low productivity workers. With O_i denoting measured output, define:

$$\text{PRODHI}_i \equiv (O_i - B)d_i$$

$$\text{PRODLO}_i \equiv (B - O_i)(1 - d_i) \ , \tag{7}$$

where $d_i = 1$ if $O_i > B$ and zero otherwise. Then, the alternative wage (index) is given from (6)-(7) as:

$$\operatorname{Ln} W_{iA} = \operatorname{Ln} W_i^* + \gamma_1^* \text{PRODHI}_i + \gamma_2^* \text{PRODLO}_i \ , \tag{8}$$

where γ_1^* and γ_2^* are unknown parameters, which were estimated as part of the complete model.

Current real wages were essentially constant over the workers in our sample. Therefore, replacing β_f with $\beta_f^* = \beta_1 \operatorname{Ln} \left[\dfrac{1 + \alpha_A}{W_{iC}(1 + \alpha_C)} \right]$ and substituting (8) into (4), we have:

$$Q_i = 1 \text{ iff } f(a_i, \delta)\,[\beta_f^* + \beta_1 \operatorname{Ln} W_{iA} + \beta_N\, \Delta N_i] > M_i \ . \tag{9}$$

Turn now to the difference in idiosyncratic benefits, ΔN_i. First, we postulate that ΔN_i depends (linearly) on PRODHI_i and PRODLO_i. Poor performers are pressured by their supervisors to work faster, while high output workers may be pressured by co-workers to shirk so that the managers will not tighten the rates. The other factors which seem important and were

included are job complexity, and the match between a worker's education level and his assigned job, and education (better educated workers may not like the simpler production jobs in our study). Match was defined as: [job complexity minus mean (job complexity in the sample)] $\times$ [education minus mean (education level in the sample]. Thus the match term captures the interaction of job complexity and education.

Therefore, we postulate that ΔN_i is given as:

$$\Delta N_i = [\text{PRODHI}_i, \text{PRODLO}_i, \text{Education}, \text{Match}_i, \text{Complexity}_i] \cdot \Theta_N , \qquad (10)$$

where Θ_N is an unknown parameter vector.

Finally, we let the costs of changing jobs depend on a constant term, a dummy variable $MF_i = 1$ for a married female and $MF_i = 0$ otherwise, and a normally distributed error term u_Q:

$$M_i = \gamma_0 + \gamma_1 MF_i + u_Q . \qquad (11)$$

With respect to MF_i, quits may be induced by a geographic change in the husbands' job as well as maternity; u_Q captures unobserved worker characteristics and random events that affect quit behavior. The quit equation is now given by substituting (10) and (11) into (9).

Before proceeding, it should be remarked that in none of the models we estimated were the data "rich" enough to identify the constant β_f^* in equation (9). To illustrate this point, consider two extreme cases that we estimated. In the first case, $\beta_f^* = 0$, which with $W_{iC}^R \approx 1.3$ corresponds to $(1 + \alpha_A)/(1 + \alpha_B) \approx 1.3$. To interpret this quantity, suppose real wages exclusive of fringe benefits were the same in the two jobs. Then, total real wages, including nonidiosyncratic fringe benefits, would be thirty percent higher in the alternative job. In the second case, we set β_f^* so that total real wages (as interpreted above) on the current job would be about two and a half times that of the alternative. Over this wide range of parameter values for β_f^*, all the estimated parameter values of the model were essentially unchanged.[5] Accordingly, we only report results below for the case in which $\beta_f^* = 0$.

3.2 Output

We let Ln(output) depend linearly on experience, age, Ln(education), dummy variables for black/white and male/female, scores on the pins and screws sections of the dexterity test, a dummy variable for workers hired in 1977, job complexity, the job matching variable defined above, and an error term, u_O.

As discussed earlier, output was normalized by "expected" output. The industrial engineering staff at the plant we examined performed the normalizations. The purpose of this normalization was to adjust output for job complexity. As argued in section 2, the industrial engineers did obtain reasonable estimates of the difficulty in performing various jobs once the requisite skills were learned. However, as also argued in section 2, their estimates of the difficulty in learning complex jobs may be biased. Accordingly, we corrected for this potential problem by including job complexity as an exogenous variable in the output equation.[6]

3.3 Education

We specify an education equation in which Ln(education) depends linearly on interactions among age, race and sex and on an error term, u_E. We have included age interactions because differences in education by race and sex may have been more prevalent in the past. To help the reader interpret our results, we included all racial and ethnic types as independent variables and omitted the intercept (constant) term (which would not now be identified).

The error term in the education equation reflects various unobservables including a "stick-to-itiveness" variable reflecting the individual's "persistence" to stick it out in school. The correlation between this error term and that in the quit equation has an interesting interpretation. Better educated workers have better alternative opportunities and may have a distaste for semi-skilled production work. Accordingly, we might expect such workers to have high quit propensities. On the other hand, high "stick-to-itiveness," which leads to increased education, may cause the worker to stay on the job longer and consequently have a low quit

propensity. To a significant extent we are able to separate out these conflicting factors by measuring the direct effect of education on quit propensity separately from the correlation of the error term in the quit equation with the error term in the education equation.

3.4 Wages

Finally, we specify a wage equation, which explains the logarithm of a worker's previous wage (recall that wages on the current job are essentially constant over workers). This variable is assumed to depend linearly on tenure, experience, education, black/white and male/female dummy variables, scores on pins and screws dexterity tests, a dummy variable for marital status, and a dummy variable for the 1977 cohort. In addition, we specify that Ln (previous wage) is a quadratic function of age, with age, a_i, entering as:

$$\theta_1(a_i{-}16) + \theta_2(a_i{-}16)^2 \ . \tag{12}$$

Here, age 16 is a normalization that does not affect the results (any number here simply amounts to redefining the constant term in the equation). We then interpret θ_1 as the impact of age on previous wage evaluated at age 16. Previous studies of wage functions lead us to expect $\theta_1 > 0$ and $\theta_2 < 0$ (for the estimates we obtain from our sample age had a positive effect on pay for all relevant ages).

As to treating tenure in the wage equation as exogenous, which is the only equation in which tenure enters (recall that quits are defined as quits within the first six months on the job), we could not reject this maintained hypothesis using a Hausman specification test.

3.5 Sample Selection Bias

From the discussion above, we have a model of the following form for individual i:

$$Q = \begin{cases} 1 \text{ if } f(a,\,\delta)[\beta_1(LnW - X_{2W}\beta_{2W}) + \beta_{1Q}^* X_{1Q}^*] - X_{2Q}\beta_{2Q} + u_Q > 0 \\ \\ 0 \text{ otherwise} \end{cases} \tag{13}$$

$$LnO = X_O\beta_O + u_O$$

$$LnE = X_E\beta_E + u_E$$

$$LnW = X_{1W}\beta_{1W} + X_{2W}\beta_{2W} + u_W \equiv X_W\beta_W + u_W \,,$$

where Q denotes quit, O denotes output, E denotes education, and W denotes previous wage. The matrices $X_W \equiv [X_{1W} \,\vdots\, X_{2W}]$, X_{1Q}^*, X_{2Q}, and X_E consist of the exogenous and endogenous variables entering the model as described above. Further, we assume that the error terms are distributed independently over individuals and for a given individual as the multivariate normal density:

$$u \equiv \begin{bmatrix} u_Q \\ u_O \\ u_E \\ u_W \end{bmatrix} \sim N(0,\,\Sigma) \,. \tag{14}$$

In estimating this model, we estimate the parameters in (13) and the covariance matrix Σ. The estimation is described in the next section.

Unfortunately, in estimating these parameters potentially serious problems of sample selection bias arise. For our sample, there are two sources of selection bias. A worker applies to work for the firm if and only if:

$$A(z) > \epsilon_1 \,, \tag{15}$$

where z is an exogenous vector of observed characteristics of the individual and of his job status, and ϵ_1 is an error term consisting of both "noise" and unobserved worker characteristics. An applicant is hired if and only if:

$$H(z) > \epsilon_2 \mid A(z) > \epsilon_1 \,, \tag{16}$$

where ϵ_2 is an error term whose interpretation is similar to that for ϵ_1. If these selection errors are correlated with errors in the rest of the model, then selection bias may result from estimating the model without taking into account the selection conditions.

Turn first to the hiring selection condition in (16). At the establishment we studied, when a job opening occurred, application forms were reviewed in the order in which they had been received. (Individuals had waited on line to submit their applications and a number was stamped on each application identifying when it was submitted. After six months, applications that had not been reviewed were discarded.) Of the applicants whose application forms were reviewed, 78% were hired. Of the rejected applicants, 85% were rejected because they did not pass the physical dexterity test. Consequently, at most 3.3% (.22 $\times$ .15) of the applicants could have been rejected on the basis of characteristics correlated with disturbances in the model. Therefore, it would appear reasonable to ignore any sample selection biases arising from the firm's hiring criteria.

Potentially more serious problems arise from sample selection biases which are due to the applicants not being a random sample from the population: the selection condition in (15). Because education, tenure and experience are positively correlated with earnings, and because some older workers have special disincentives for changing jobs,[7] those better educated and more experienced workers who apply for these jobs are likely to have unobserved characteristics which cause them to have low values of ϵ_1. Accordingly, such workers are more likely to change jobs or to leave unemployment for a semi-skilled manufacturing job. Fortunately these biases are mitigated by the high wages paid by the establishment we studied. The firm was unionized, and the median payment of these new hires (who were employed elsewhere when they applied for this firm) was 203% of their median previous pay. These high wages lead to an excess supply of applicants; some applicants waited in line, in freezing temperatures, for over 36 hours to receive application forms. By paying wages above the market wage for broad categories of workers, the firm is mitigating the sample selection bias problem. To see this, suppose the firm were to pay

an infinite wage. Then, all workers would apply to the firm, and if the order of applications were random, there would be no sample selection problems arising from the application process. Therefore, since we have reason to believe that the sample selection bias is small and since an attempt to "correct" for it would be based on strong and tenuous distributional assumptions, we have chosen not to correct for this bias.[8]

As a further check on the importance of sample selection bias, we compared the returns to education for the workers in our sample on their previous job with returns to education estimated in the Current Population Survey. If our sample is significantly biased we would expect to find exceptionally low returns to education for workers in our sample. The better educated workers in our sample would be those who did exceptionally poorly in the labor market. This was not the case. Mellows found that in 1979 the rate of return to education for the 18,551 workers in the Current Population Survey was between 4% and 7%, depending on the independent variables he included. For the workers in our sample, our maximum likelihood estimates of their rate of return to education on their previous job was 5-6%. Thus, for our sample, education made roughly the same (or a slightly greater) contribution to their wages on their previous job than in the larger random CPS sample. Because of our relatively small numbers of degrees of freedom and the limited range of education levels in our sample, our estimates of rates of return to education are very imprecise. However, the similarity of the estimated rate of return to education between the two samples is consistent with a lack of serious sample selection bias in our sample.

4. THE ESTIMATION METHOD

Since the model in (13)-(14) is estimated by maximum likelihood, in this section we formulate and discuss the likelihood function. From (13)-(14), let q_i be the quit probability for individual i conditioned on output, education, and wages:

$$q_i \equiv \mathbf{P}\left\{f(a,\,\delta)\,[X_{1Q}\,\beta_{1Q}] - X_{2Q}\,\beta_{2Q} + u_{Qi} > 0 | \mathrm{Ln}\,O_i,\; \mathrm{Ln}E_i,\; \mathrm{Ln}\,W_i\right\}. \qquad (17)$$

Next, let g_i be the joint density pertaining to individual i for $(\mathrm{Ln}\,O_i,\; \mathrm{Ln}E_i,\; \mathrm{Ln}\,W_i)$:

$$g_i \equiv g_i(\mathrm{Ln}\,O_i,\; \mathrm{Ln}E_i,\; \mathrm{Ln}\,W_i)\,. \qquad (18)$$

In both (17) and (18), it is implicit that we always condition on all exogenous variables in the model. Then, with Q_i as the quit random variable, equaling one for a quit and zero otherwise, the log likelihood is given by:

$$\mathrm{Ln}L = \sum_{i=1}^{N}\left[Q_i\,\mathrm{Ln}\,q_i + (1 - Q_i)\,\mathrm{Ln}(1 - q_i)\right] + \sum_{i=1}^{N}\mathrm{Ln}\,g_i\,. \qquad (19)$$

Notice that we have assumed that we have independent observations on the $i = 1,...,N$ individuals.

To completely specify the likelihood function, it now remains to formulate the conditional quit probability, q_i, and the joint density, g_i. Beginning with the conditional quit probability, note that as in all probit models, the variance of the quit disturbance, u_Q, is not identified. Accordingly, we normalize this variance to 1. Then, we are estimating parameters of the quit equation relative to this variance. Consequently, from (14), we have:

$$u_i \equiv \begin{pmatrix} u_{Q_i} \\ u_{E_i} \\ u_{O_i} \\ u_{W_i} \end{pmatrix} \sim N(0,\,\Sigma)\,, \qquad (20)$$

$$\Sigma \equiv \begin{bmatrix} 1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

and

$$u_{Ri} \equiv \begin{pmatrix} u_{Ei} \\ u_{Oi} \\ u_{Wi} \end{pmatrix} \sim N(0, \Sigma_{22}) \ .$$

Therefore, the distribution for u_{Qi} given u_{Ri} is:

$$u_{Qi} | u_{Ri} \sim N\left(\Sigma_{12} \Sigma_{22}^{-1} u_{Ri}, \ 1 - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right) . \tag{21}$$

We can now employ (21) to derive the desired conditional quit probability, q_i. To this end, from (13) we have:

$$u_R = \begin{pmatrix} LnE_i \\ LnO_i \\ LnW_i \end{pmatrix} - \begin{pmatrix} X_{Ei} & 0 & 0 \\ 0 & X_{Oi} & 0 \\ 0 & 0 & X_{Wi} \end{pmatrix} \begin{pmatrix} \beta_E \\ \beta_0 \\ \beta_W \end{pmatrix} \equiv Y_{Ri} - X_{Ri}\beta_R \ . \tag{22}$$

Let u_{Qi}^* be a random variable with distribution:

$$u_{Qi}^* \sim N(\Sigma_{12}\Sigma_{22}^{-1}(Y_{Ri} - X_{Ri}\beta_R), \ 1 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) \ . \tag{23}$$

Then, from (13) and (21) the conditional quit probability in (17) is now given by:

$$q_i = \mathbf{P}(f(a, \delta)[X_{1Q}\beta_{2Q}] - X_2\beta_{2Q} + u_Q^* > 0) \tag{24}$$

$$= \Phi\left\{ [f(a, \delta)[X_{1Q}\beta_{2Q}] - X_2\beta_{2Q} + \Sigma_{12}\Sigma_{22}^{-1}(Y_{Ri} - X_{Ri}\beta_R)] / [1 - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}]^{1/2} \right\},$$

where Φ is the cumulative distribution function for a random variable distributed as $N(0, 1)$.

Turning now to g_i, the joint density for the other endogenous variables in the model, this density is multivariate normal by assumption. From the description of the model, except for correlations among the disturbances, the sub-model pertaining to these endogenous variables has a recursive form. Accordingly, the Jacobian determinant from the disturbances to these endogenous variables is unity. Consequently, from (13), (14), and (22):

$$g_i(LnE_i, LnO_i, LnW_i) \equiv g_i(Y_{Ri}) = N(X_{Ri}\beta_R, \Sigma_{22}) \ . \tag{25}$$

The likelihood is now defined by substituting (24) and (25) into (19).

Before proceeding to discuss the results of maximum-likelihood estimation, it is important to point out properties of the estimator under both correct and incorrect model specifications. Let $\hat{\theta}$ be the value of the parameter vector θ that maximizes the likelihood given above, denote N as the sample size, and let θ_0 be the true value of the parameter vector. Then, if the model is correctly specified, the asymptotic distribution of $\sqrt{N} \ (\hat{\theta} - \theta_0)$ is:

$$\sqrt{N} \ (\hat{\theta} - \theta_0) \rightarrow N(0, \ \Omega) \ . \tag{26}$$

With $L_i(\theta)$ as the log likelihood, for the ith observation and $Q_N(\theta)$ as the sample averaged log likelihood, for a correct mode we can estimate Ω consistently by:

either

$$\hat{\Omega}_H \equiv - \ (\partial^2 Q_N(\hat{\theta}) / \partial\theta \ \partial\theta')^{-1} \tag{27}$$

or

$$\hat{\Omega}_G \equiv \left\{ (1/N) \ \sum_{i=1}^{N} \left[(\partial L_i(\theta) / \partial\theta)(\partial L_i / \partial\theta)' \right] \right\}^{-1} , \tag{28}$$

where with $\theta : p \times 1$, $\partial L_i / \partial\theta : p \times 1$. We refer to $\hat{\Omega}_H$ as the Hessian estimate and $\hat{\Omega}_G$ as the Gradient estimate.

When the model is misspecified, denote by "quasi-likelihood" the hypothesized as opposed to the true likelihood. From White's (a) results, we can interpret the quasi-likelihood estimate, $\hat{\theta}$, as consistently estimating the parameter value, θ^*, that minimizes the Kullback-Leibler distance between the quasi-likelihood and the true likelihood. Furthermore, under conditions given by White (a), we have the asymptotic result:

$$\sqrt{N} \ (\hat{\theta} - \theta^*) \rightarrow N(0, \ \Omega^*) \ . \tag{29}$$

Under conditions given by White (a,b), we can estimate Ω^* consistently by:[9]

$$\hat{\Omega}_W = \hat{\Omega}_H \ \hat{\Omega}_G^{-1} \ \hat{\Omega}_H \ . \tag{30}$$

In reporting our results below, we will report standard errors for $\hat{\theta}$ obtained from Hessian, Gradient, and White's estimated covariance matrices. If the model has been specified incorrectly, we would expect considerable divergence between these estimated standard errors. In the next section, we consider several specification tests, including a comparison of Hessian and Gradient covariance matrices. Then we turn to the main results of our study.

5. MODEL SPECIFICATION TESTS

In this section, we present our specification results for workers (463) who were employed at the time they applied for their current job. From previous discussions, our model is not designed to explain the behavior of workers who were unemployed when they applied for their current job. For example, adjusted previous pay is most indicative of alternative opportunities for workers who were employed at time of hire. Moreover, as discussed earlier, the presence of such workers results in serious specification problems, which is reflected in a large divergence in the covariance matrix estimates discussed above [i.e., in (27), (28), and (30)].

In interpreting our results below, we will be concerned in part with whether there is a factor, termed "stick-to-itiveness", that is associated with individuals remaining in both school and on the job. To the extent that stick-to-itiveness is not learned in secondary school, the correlation between quit and education errors should be negative and significant. By estimating this correlation, we allow for this possibility. Although it seems unlikely to us that stick-to-itiveness is learned in secondary school, we did examine this possibility. We allowed for the possibility that stick-to-itiveness is learned in secondary school by including observed education as an element of moving costs (i.e., X_{2Q} in (17)): that is, education entered the quit equations both interacted with $f(a, \delta)$ and without the interaction effect. If stick-to-itiveness is learned in secondary school, then the coefficient on this education variable (without the interaction with $f(a, \delta)$) should be negative and statistically significant. We enter the education variable here in log form, because this is the metric comparable to the error term in the education equation

(though, alternative specifications yielded similar results).

In Model A1, shown in Table A1 of the Appendix, we address the above specification issue. As an element of X_{2Q}, education has a negative coefficient (-.04) that is not statistically different from zero (its standard error exceeds 7.26). Moreover, we obtained this result for a variety of other specifications in which education enters as an explanation of moving costs. Consequently, it would appear that stick-to-itiveness is not learned in secondary school.

One could argue that the above finding might be an artifact of multicollinearity. With education entering alternative opportunities, moving costs, and the error structure, it may be difficult to disentangle these effects. In particular, from Table A1, education is not a statistically significant determinant of alternative opportunities or of moving costs. Consequently, it is indeed difficult to disentangle these two effects. Nevertheless, the correlation between the error terms in the quit and education equations remains negative (-.76) and highly significant in Table A1 (this is our estimate of stick-to-itiveness). Therefore, while we cannot eliminate the possibility that some component of "stick-to-itiveness" is learned in secondary school, there does appear to be an important component that is not learned in secondary school. In what follows, we will report results and conduct specification tests, for the model shown in Table 2, where "stick-to-itiveness" is not learned in secondary school.

In examining the model in Table 2, Hessian, Gradient, and White standard errors are similar: inferences about statistical significance are not sensitive to which of these standard errors is employed. Recall that if the model is specified correctly, there should be little divergence between these estimated standard errors. Further, recall that this divergence was considerable when we included workers who were unemployed at time of hire. Such workers were consequently removed from the sample.

As a formal test of the model's specification, we could perform White's (a,b) information test. This test is based on a quadratic measure of the distance between the gradient outer products

and the second partial derivatives of the log likelihood function. Unfortunately, it is computationally infeasible to implement this test for the entire model, as this test would require us to invert a matrix of dimension 1378.[10] Accordingly, we performed this test on parameter subsets. The results depended on the number of parameters in the subset. Even neglecting this ambiguity, the results of these tests are difficult to interpret. As White (a) notes, failure to pass these tests, may indicate either inconsistent parameter estimates, or consistent parameter estimates but incorrect estimates of the standard errors being generated by the Hessian or Gradient outer product matrices. In the latter case, the appropriate covariance matrix would be that given by White (a). It may also be the case that the sample size here is too small for the asymptotic distribution of the test statistic to hold (approximately).

To test for inconsistency in parameter estimates, White proposed a Hausman type test of the model, which we now examine. Here, we focus on the quit equation's specification for two reasons. First, several of our most important findings pertain to the quit equation. Second, because quits are described by a probit equation and all other equations are linear, our results are most sensitive to misspecification in the quit equation.

Let θ_1 be the parameters in the submodel not containing the quit equation. Let $\hat{\theta}_1^*$ be the estimate of θ_1, obtained by maximizing the "postulated likelihood" for the entire model including the quit equation. Let $\hat{\theta}_1$ be the estimate of θ_1, obtained by maximizing the "postulated likelihood" for the sub-model excluding the quit equation. Notice that unlike $\hat{\theta}_1^*$, the consistency of $\hat{\theta}_1$ does not depend on whether or not the quit equation is specified correctly. Then, with N as the sample size and V as the asymptotic covariance matrix of $\sqrt{N}(\hat{\theta}_1^* - \hat{\theta}_1)$, the test statistic is:

$$N(\hat{\theta}_1 - \hat{\theta}_1^*)' \ V^{-1}(\hat{\theta}_1 - \hat{\theta}_1^*) \ . \tag{31}$$

Under general conditions this test statistic is asymptotically distributed under the null hypothesis as is χ^2 with p_1, the dimension of θ_1, degrees of freedom.[11] The problem now is to estimate the covariance matrix V. Suppose that under the null hypothesis not only are $\hat{\theta}_1$ and $\hat{\theta}^*$

consistent, but that the entire model is specified correctly. Then from Hausman, V can be estimated consistently by the difference in consistent estimates of the covariance matrices for $\hat{\theta}_1$ and $\hat{\theta}_1^*$ respectively. (Note that such an estimate need not be positive semi-definite [see footnote 12 below].) Since our null hypothesis is that $\hat{\theta}_1^*$ and $\hat{\theta}_1$ are consistent, we propose to estimate V by a White-type estimator.

Let θ_0 be the true value of θ, θ_{10} the true value of θ_1, L_1 the average log likelihood for the sub-model, and g_1 the gradient of L_1. Let L^* and g^* be the corresponding quantities for the full model. Then, with H^* as the appropriate partition of $[\partial^2 L^*(\theta_0)/\partial\theta\,\partial\theta']^{-1}$, we have:

$$\sqrt{N}(\hat{\theta}_1^* - \hat{\theta}_1) \approx H^*(\theta_0)\,\sqrt{N}\,g^* - H_1(\theta_{10})\sqrt{N}\,g_1\,, \tag{32}$$

where N is the sample size,

$$H_1(\theta_{10}) \equiv [\partial^2 L_1(\theta_{10})/\partial\theta_1\,\partial\theta_1']^{-1}\,, \tag{33}$$

and $\approx$ means equal in asymptotic distribution. Therefore, the asymptotic covariance matrix of $\sqrt{N}(\hat{\theta}_1^* - \hat{\theta}_1)$ is

$$V = E\left\{ |H^*(\theta_0)\,\sqrt{N}\,g^* - H_1(\theta_{10})\,\sqrt{N}\,g_1| \right\}, \tag{34}$$

where $|A| \equiv AA'$. Under conditions given by White (a), we can then estimate V consistently by:[12]

$$\hat{V} = (1/N) \sum_{i=1}^{N} |\Omega_i|\,, \tag{35}$$

$$\Omega_i \equiv H^*(\hat{\theta}^*)\sqrt{N}\,g_i^*(\hat{\theta}^*) - H_1(\hat{\theta}_1)\,\sqrt{N}\,g_{1i}(\hat{\theta}_1)\,,$$

where g_i^* and g_{1i} are the indicated gradients for the ith observation evaluated at $\hat{\theta}^*$ and $\hat{\theta}_1$ respectively.

With $\hat{V}$ given as in (35), the value of the test statistic in (31) is 9.8. Since the critical χ^2 value at 39 degrees of freedom exceeds 50 at the 1% significance level, we can not reject the null

hypothesis that the quit equation is estimated consistently.

As an additional specification test, we also constructed a Box-Cox type test of the quit equation. Recall that the quit equation has the form:

$$Q = 1 \text{ if } \psi(Z_Q, \beta_Q) + u_Q > 0 \tag{36}$$

$$= 0 \text{ otherwise}$$

where u_Q is assumed to be i.i.d. $N(0, 1)$ (the variance of u_Q can be taken to be 1.0 without loss of generality). Consider now a generalization of (36) obtained by successive exponential and Box-Cox transformations. Namely,

$$Q = 1 \text{ if } \left\{ [\exp \psi(Z_Q, \beta_Q)]^\lambda - 1 \right\} / \lambda + u_Q^* > 0 \tag{37}$$

$$= 0 \text{ otherwise,}$$

where now u_Q^* is assumed to be $N(0, 1)$. As $\lambda \to 0$,

$$\left\{ [\exp \psi(Z_Q, \beta_Q)]^\lambda - 1 \right\} / \lambda \to \text{Ln } e^{\psi(Z_Q \beta_Q)} = \psi(Z_Q \beta_Q) \ . \tag{38}$$

Therefore, assuming quit behavior is described by the generalization in (38), we can test the specification in (36) by testing (37) with $\lambda = 0$ as the null hypothesis.

To perform this test we re-estimated the entire model with the generalized quit equation as given in (37). Since parameter estimates of the entire model are not significantly affected by generalizing the quit equation, Table A2 of the Appendix reports results only for the generalized quit equation. From Table A2, λ is estimated to be .18 with a standard error exceeding .14. Consequently, we can not reject the null hypothesis that $\lambda = 0$. Perhaps not surprisingly, given this finding, if we compare estimates of the quit equations in (36) and (37), from Tables 2 and A2 there is no great change in any of the parameter estimates. Since all of our inferences below hold for both quit equations and since we cannot reject $\lambda = 0$, we only discuss full results for the case

in which $\lambda = 0$. These results are reported in Table 2.

Although this section has focused on formal model specification tests, perhaps the most persuasive evidence that the model in Table 2 is not grossly misspecified is that the maximum likelihood value of δ was .07 (the net discount rate). Since we placed no constraints on the values δ could take, it was reassuring that the best fit was obtained with such a plausible value.

To summarize this discussion, we do not contend that the model is exactly correct. Almost surely, all economic models are misspecified. Moreover, we have not considered all possible specification tests, but rather those that seem the most relevant. These tests suggest that the model in Table 2 is sufficiently well specified to warrant an examination of its implications.

6. RESULTS

6.1 Education

The first results we shall discuss concern explanations for the higher wages, on their previous jobs, of the better educated workers in our sample. There is considerable disagreement over the role of education in a standard wage equation. The positive effect of education may be a consequence of learning in school which increases worker productivity. Alternatively, it may be that individuals who go to school longer also have unobserved traits (such as stick-to-itiveness) which make them more productive.

To explain the positive relationship between education and previous wages, we first looked at the output of better educated workers. The output equations estimated in Table 2 suggests that output on these semi-skilled production jobs is not correlated with education, nor do better educated workers perform significantly better on more complex tasks — although the sign of the match term is positive, the estimated effect on output is not statistically or economically significant. From Table 2, we estimate that a one standard deviation increase in the match term leads to a .02 increase in output. Since the mean output is 111 this increase is trivial.

One particularly striking result of the model is the large negative and statistically significant correlation of the errors in the quit and education equations. For the model estimated in Table 2, which the specification tests in section 5 suggest is the preferred model, this correlation is -.76 and (in every model we estimated) is statistically significant: Individuals who for unexplained reasons stick-it-out on the job are those who choose more education than would be expected from their observed characteristics. As discussed in section 5 (Specification Tests) and shown in Tables 2, A1, and A2, this effect persists under a variety of alternative specifications of the quit equation.

When we estimated quasi-reduced form quit equations in which the independent variables being multiplied by $f(a, \delta)$ were: race, sex, age, education, scores on each section of the dexterity test, match, job complexity and a constant term, the estimated correlation between the error term of the quit equation and the error term on the education equation ranged from -.90 to -.95, depending on whether output and marital status were included and the interactions allowed among the independent variables. Hence the negative correlation of the errors in the quit and education equations appears to be a robust result for our sample.

Extensive conversations with personnel officers indicate that a low quit propensity is one of the traits most valued in employees. Therefore, it may be that one of the unobserved attributes that play an important role in explaining the correlation between education and earnings is the lower quit propensity of better educated workers.

Education also affects the quit propensity directly through its effect on job satisfaction. As expected, individuals with more education are more likely to have lower levels of job satisfaction on these jobs. Such individuals are more likely to quit, the relevant estimated education coefficient here is .053 from Table 2, with estimated standard errors ranging from .020 to .024. (Recall that we already included their superior alternative opportunities in the adjusted previous pay variable in the quit equation.)

The higher quits of the better educated workers (controlling for adjusted previous pay) could be explained differently if the better educated workers who took these jobs were relatively underpaid on their previous jobs — had had worse draws from their wage distribution. In that case, adjusted previous pay would understate the expected value of the alternative opportunities of the better educated workers. This would cause a positive partial correlation between quite rates and education.

However, our estimate of the rate of return to education on their previous jobs for workers in our sample was similar to estimates of this return for workers sampled by the Current Populations Survey during the same period (see section 3.5). Consequently, it does not appear that the better educated workers were relatively underpaid on their previous job. The positive correlation between quit rates and education seems to be better explained by differences in job satisfaction than by biases in the relationship between adjusted previous pay and the best alternative wage offer.

6.2 Male/Female and Black/White Wage Differences

Another issue we were able to investigate is whether direct measures of output can help explain male/female or white/nonwhite pay differentials. We estimate that being a male increases a worker's earnings on his previous job by 22%, while the output of men on their current jobs exceeds the output of women by 9%. Similarly, we estimate that being white increases earnings on the previous job by 5.5%, and that first month output on their current jobs for whites is 2.5% higher than for nonwhites (for whites and non-whites neither pay nor output differences are statistically significant). We offer three explanations for the disparity between pay and output differences.

(1) If previous employers had significant costs per worker, a given percentage difference in output per hour may result in a much larger difference in the value of a worker to a firm. For example, assume that some of the overhead, capital costs, and benefit payments are a function of

the number of hours worked. Further assume that these costs per hour worked equal the average wage received by a male worker. Then, the firm's cost of the hourly output of a male worker is $2w_m + p$, where w_m is the male wage and p is the cost of materials. If males are 9% more productive than females, the cost of the same output produced by a female worker is 1.09 $(w_f + w_m) + p$, where w_f is the female wage. If both males and females are employed, (absent discrimination) the cost of producing goods using male workers must be the same as the cost using female workers:

$$1.09(w_f + w_m) + p = 2w_m + p, \quad \text{or}$$

$$w_m/w_f = 1.20 .$$

Therefore, a 9% output differential would translate into a 20% wage differential. Since we find that the wage differential is 22%, the observed output difference, coupled with a reasonable estimate of costs that are proportional to the number of hours worked, could account for the male/female wage differential. Of course, we would need to have more evidence on the proportion of costs that are independent of output but a function of total hours worked to answer this question adequately. Moreover, if women learn more rapidly on the job than do men, then the 9% output differential (output, for reasons given earlier, is measured in the first month of the job) may not persist.

(2) The most obvious "explanation" is that quit propensity and output on the current job may be a poor indicators of performance on previous jobs. In this case, we would have insufficient information to determine whether or not discrimination existed on the previous job.

(3) Absent significant per employee costs or omitted aspects of performance that are correlated with race or sex, the difference between wage and output differentials suggests the presence of wage discrimination on previous jobs. To provide one explanation as to why such discrimination could persist, suppose that previous employers (incorrectly) perceive that women have significantly higher quit propensities than men, and that Viscusi is correct in asserting that

"the presence of substantial [perceived] differences in quit behavior may account for substantial pay differences between men and women."[13] Next, suppose through some historical accident, women are overrepresented in jobs with relatively flat wage profiles. Then, even if males and females have the same quit propensities (holding alternative opportunities the same for males and females), differences in job assignment would cause females to have higher quit rates. Employers "learn" that females have high quit rates and mistakenly conclude that it is unprofitable to train women as much as men. If women receive less training and thus acquire less firm specific capital than men, they will be observed in jobs with relatively flat wage profiles and high quit rates. Therefore, on the previous jobs for the workers in our sample, wage discrimination could persist. (Recall that the wage equation explains previous pay; men and women have identical current wage profiles at the firm we are examining.)

6.3 Alternative Opportunities

As we've mentioned, the model estimated in Table 2 places considerable economic structure on the behavior of individuals. In particular, individuals quit these jobs if and only if their alternative opportunities exceed the value they place on their current job by an amount sufficient to outweigh their pecuniary and non-pecuniary costs of changing jobs. As the model predicts, $f(a, \delta) \times$ adjusted previous wage, has a positive effect on quits: individuals whom our model predicts have better alternative opportunities are also more likely to quit (this effect varies from being statistically insignificant in Table 2, to being of border line significance at the 10% level in Table 2A).

Similarly we would expect that individuals with exceptionally high output would also have better than average alternative opportunities. On the other hand, individuals with low output might be induced to quit by pressures from their supervisors. Again the predictions of the model are consistent with the estimates we obtained. Individuals with the highest (PRODHI) output levels are both more likely to quit these jobs than are individuals whose output is near the median. The effect of being a low productive worker (PRODLO) is not significant in this model.

The estimated coefficients in the quit equation thus appear consistent with the predictions of the model formulated in section 2.

6.4 Age vs. Experience

We estimate that at the ages of majority of the workers in our sample, the partial correlation between age and previous wages is positive and statistically significant while the partial correlation between previous experience and previous wage is not statistically significant.[14] One explanation for this phenomenon is that the measure of experience, which is computed from the reported starting and ending dates of the applicants' previous jobs, is so faulty that age is a better proxy for true experience than is reported experience.[15] However, our sample is quite young; it seems likely that most of the applicants would have a good recollection of how long they held their previous jobs. Finally, the tenure coefficient is positive and has approximately the same effect as Mellow(a,b) estimated for the C.P.S. sample. Thus it appears likely that respondents' reports of tenure on their immediately previous job are reasonably accurate. If the recorded levels of experience are accurate, then the significant effect of age on earnings and the insignificant effect of experience can be explained by:

(1) Older workers having lower quit propensities, as can be seen from the coefficients in the estimated quit equation;

(2) Older workers having lower rates of absenteeism. (Weiss [1985a] shows that, holding other demographic characteristics fixed, older production workers at this firm are absent significantly less often. Weiss [1985b] presents conditions under which absenteeism can have a disproportionate effect on wage.)

(3) The longer individuals search for jobs the better the job they are likely to have. Since individuals engage in search both while employed and unemployed, age may better approximate total time spent searching than does experience: age measuring the amount of job search an individual has engaged in. (In many wage studies the

proxy used for experience is age minus education minus five which may be a better approximation of workers' ages than of their true work experience.)

6.5 Job Matching

We argued earlier that we might expect that assigning the better educated worker to the more complex jobs would affect both their output and their quit propensity.[16] This expectation could not be confirmed in the data. The match term was statistically insignificant in both the output and quit equations. (In the quit equation it also had the "wrong" sign, however, our estimates of the standard errors were all more than four times as large as the estimated coefficient.) Consequently, we were not able to conclude that better educated workers had a comparative advantage on relatively complex semi-skilled production jobs — nor could we find any evidence that they prefer those jobs more (dislike them less) than do the less well educated workers. Not surprisingly, as discussed in section 6.1, better educated workers appear to have a distaste for all of the jobs at the plant we studied.

FOOTNOTES

1. In a sophisticated human capital model the choice of an education level is determined (in part) by the efficiency with which an individual learns. In those models, as in sorting models, unobserved attributes that effect the cost (or benefit) from schooling are correlated with levels of education.

2. A literal transcription of the recorded data in the Panel Study of Income Dynamics (PSID) would find that women age more slowly than men.

3. If the model is specified correctly, then standard errors calculated from the log likelihood Gradient should be equivalent to those calculated from the Hessian (this equivalence forms the basis for White's information test). In several instances, such standard errors were radically different. For example, when the sample included workers who were unemployed when they were hired, the standard errors differed by an order of magnitude greater than 10.

4. All the data used in this study were collected at the plant before the study was contemplated.

5. By way of a summary measure, the log-likelihood ranged from 358.6 to 358.8 for the values of β_f^* we imposed.

6. With O^* as the output variable in the output equation and γ the coefficient on job complexity, JC, in this equation, correct output would be $O^* - \gamma \cdot JC$. Since we found γ positive and statistically significant, it would appear that O^* is biased upwards.

7. These disincentives include pension plans which are convex functions of tenure, and firmer ties with coworkers as one stays on a job longer.

8. To correct for this bias, we could assume a joint distribution encompassing not only the error terms in our model, but also those in the selection equations. Then, conditioned on

$A(z) > \epsilon_1$ and $H(z) > \epsilon_2$, we could proceed to formulate and maximize a likelihood function. (This is the methodology suggested by Bloom and Killingsworth.) Our results would then depend on what could very well be a misspecification of the error distributions in the selection equations or a misspecification in the selection equations themselves.

9. We must assume here, for example, that the exogenous variables are identically and independently distributed. (See White (a,b).)

10. With p parameters in the model, there are potentially $p(p + 1)/2$ distinct Gradient outer product and Hessian terms. The average differential between these terms (averaging over the observations) yields a variable whose covariance matrix has dimension $p(p + 1)/2$. In White's test, this matrix must be inverted, which is infeasible for our model with $p = 53$.

11. From Taylor series expansions, we have:

$$\sqrt{N}\ (\hat{\theta}_1^* - \theta_{10}) \approx \mathbf{P}\ \left\{\partial^2 L^*(\theta_0)/\partial\theta\,\partial\theta'\,^{-1}\right\}\sqrt{N}\ g_*(\theta_o)$$

$$\sqrt{N}\ (\hat{\theta}_1 - \theta_{10}) \approx [\partial^2 L_1(\theta_{10})/\partial\theta_1\,\partial\theta_1'\,]^{-1}\sqrt{N}\ g_1(\theta_{10})$$

where $\mathbf{P}$ is a partition operator that chooses the appropriate partition of $[\partial^2 L^*/\partial\theta\,\partial\theta']^{-1}$, and where $\approx$ means equal in asymptotic distribution. From the difference in the above expressions, we find an expression for $\sqrt{N}\ (\hat{\theta}_1^* - \hat{\theta}_1)$. We can re-write this expression as a vector, each of whose elements is the sum of N independent terms. Each term corresponds to an observation and consists of a linear combination of gradients. Formally, then, the argument requires us to apply a central limit theorem to this sum of independent terms.

12. In particular, let X_i be an observation on the exogenous variable vector. Then, we require that this vector be independently drawn from a common distribution. Subject to other conditions given in White (a), the covariance matrix estimate $\hat{V}$ in (35) will be consistent. If we did assume, which we do not, that the entire model is specified correctly, then the covariance matrix can be estimated by the difference in covariance matrices for $\hat{\theta}_1^*$ and $\hat{\theta}_1$.

Notice that we require more than the consistency of $\hat{\theta}_1^{'}$ and $\hat{\theta}_1$ for this calculation to be valid. We did indeed calculate this difference, which turned out to be an indefinite matrix. The consistent estimate $\hat{V}$ in (35) will always be positive semi-definite and can be expected to be positive definite.

13. Viscusi, p. 389. Note also that Flinn shows that of individuals who have had at least two job changes within their first year of entering the labor force have earnings substantially below those of workers with no job changes.

14. Of course, experience is an endogenous variable. We are treating it as exogenous, which may introduce some bias into our results. However we would expect the error term in the experience equation to be positively correlated with the error term in the wage equation (the substitution effect outweighing the income effect). In that case our estimate of the coefficient on experience in the wage equation is likely to be biased upward, and our estimates would understate the disparity between the effect of experience and of age on previous wages.

15. It is also possible that the insignificance of experience and the significant effect of age on earnings is an artifact of our sample. We have argued in section 2 that sample selection bias is unlikely to have a strong effect on our results. Nevertheless, when a sample is truncated and the truncation is affected by the dependent variable (in our case very highly paid workers do not apply for the jobs we study), the estimated coefficients on the included variables are generally biased.

16. Recall that job complexity (JC) as defined in section 2, measures the training required to perform the job. The match variable was defined as: [Education — Mean Education] [JC — Mean JC].

REFERENCES

Arnold, Hugh and Daniel Feldman, "A Multivariate Analysis of the Determinants of Job Turnover," *Journal of Applied Psychology,* Vol. 67, No. 3, (June 1982) pp. 350-360.

Bartel, Ann, "Wages, Nonwage Job Characteristics and Labor Mobility," *Industrial and Labor Relations Review,* Vol. 35, No. 4 (July 1982) pp. 578-589.

Bloch, Farrell, "Labor Turnover in U. S. Manufacturing Industries," *Journal of Human Resources,* Vol. 14, No. 2 (Spring 1979) pp. 236-245.

Flinn, Christopher, "Wages and Job Mobility of Young Workers," unpublished mimeo 1983.

Garen, John, "The Returns of Schooling, A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica,* Vol. 52, No. 5 (September 1984) pp. 1199-1218.

Hausman, J. A., "Specification Tests in Econometrics," *Econometrica,* Vol. 46, No. 6 (November 1978) pp. 1251-1272.

Heckman, James, "The Common Structure of Statistical Models of Truncation. Sample Selection, and Limited Dependent Variables and a Simple Estimator for such Models," Annals of Economic and Social Measurement, Vol. 5, (1976) pp. 475-492.

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica,* Vol. 47, (1979) pp. 73-92.

Knapp, Martin; Kostas Harissis and Spyros Missiakoulis, "Investigating Labor Turnover and Wastage Using the Logit Technique," *Journal of Occupational Psychology,* Vol. 55, (1982) pp. 129-138.

Mellow, Wesley, "Employer Size and Wages" *Review of Economics and Statistics,* Vol. 64, No. 3 (August 1982) pp. 495-501.

Mobley, W. H., Griffeth, H. H. and Meglino, B. M., "Review and Conceptual Analysis of the Employee Turnover Process", *Psychological Bulletin,* (1979) Vol. 86, pp. 493-622.

Pencavel, John, "Wages, Specific Training and Labor Turnover in U. S. Manufacturing Industries," *International Economic Review,* Vol. 13, No. 1, (1972) pp. 53-64.

Polachek, Solomon, "Occupational Self-Selection: A Human Capital Approach to Sex Differences" *Review of Economics and Statistics,* Vol. 63, No. 1, (February 1981) pp. 60-69.

Steers, R. and S. Rhodes, "Major Influences on Employee Attendance: A Process Model" *Journal of Applied Psychology,* (1978) Vol. 63, No. 4, pp. 391-407.

Viscusi, Kip "Sex Differences in Worker Quitting" *Review of Economics and Statistics,* Vol. 62, No. 3 (August 1980) pp. 388-398.

Weiss, Andrew, "Determinants of Quit Behavior for Semi-skilled Production Workers", *Journal of Labor Economics,* Vol. 2, No. 3. (July, 1984) pp. 371-387.

Weiss, Andrew "Testing the Sorting Model of Education," Columbia University discussion paper No. 245, (April 1984).

Weiss, Andrew "High School Graduation, Performance and Earnings": NBER Working Paper No. 1595, (April 1985a).

Weiss, Andrew "Wages and Absenteeism'" *Economics Letters*, No. 19, (1985b) pp. 277-279.

White, Halbert "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, Vol. 50 (1982) pp. 1-25.

White, Halbert "Corrigendum," *Econometrica*, Vol. 51, No. 2 (March 1983) pp. 513.

APPENDIX: SPECIFICATION TESTS

TABLE A1: QUIT BEHAVIOR — SIGNALED BY EDUCATION vs. LEARNED IN SECONDARY SCHOOL[a]

DEPVAR=QUIT[b]	VALUE	GRADIENT SE	HESSIAN SE	WHITE SE
CONSTANT	-8.2598215	13.2666076	12.1776174	17.8263665
$f(a,\cdot) \times$ Ln ADJ. PREV. WAGE	0.0710270	0.0766275	0.0607788	0.0649694
$f(a,\cdot) \times$ PRODLO	-0.004955	0.0011376	0.0006388	0.0004603
$f(a,\cdot) \times$ PRODHI	0.0023593	0.0015874	0.0013020	0.0016811
$f(a,\cdot) \times$ EDUCATION	0.0528034	0.0758144	0.0532811	0.0697054
$f(a,\cdot) \times$ MATCH	0.0013001	0.0104776	0.0053839	0.0062142
$f(a,\cdot) \times$ JOB COMPLEXITY	-0.0049965	0.0116252	0.0072276	0.0063711
MARRIED FEMALE	0.3418052	0.2310382	0.1846040	0.1938735
Ln(EDUCATION)*	-0.0414891	8.4330375	7.2633634	10.5207627
DISCOUNT RATE (δ)	0.0734658	0.0405998	0.0290190	0.0335725
SPLIT (B)	142.4822382	16.2011333	13.3756011	12.6953582
CORRELATION: Quit, Ed. Errors	-0.7573322	0.1794186	0.1715474	0.3511814

FINAL Ln LIKELIHOOD = 358.797435 (FULL MODEL)

---

a Education enters this model in three ways. First, highly educated workers may have a distaste for their current production jobs. Such disutility should be discounted, in which case education interacts with the discount factor $f(a, \cdot)$ (see the discussion of the quit equation). Second, the correlation between education and quit errors (-.75 here) may reflect "stick-to-itiveness" not learned in secondary school. Third, workers may learn in secondary school to "stick-it-out" on the job. This effect enters as do moving costs and is captured by the Ln (Education) term. Alternative specifications, with different metrics for the variables, were also examined. Results were similar with lower likelihoods.

b See Table 2 for definitions of $f(a, \cdot)$, PRODLO, PRODHI, and B.

TABLE A2: THE NORMALITY ASSUMPTION IN THE QUIT EQUATION

	VALUE	GRADIENT SE	HESSIAN SE	WHITE SE
DEPVAR = QUITS[a]				
CONSTANT	-7.762	2.828	1.553	1.490
$f(a,\cdot) \times$ ADJ. PREV. PAY	0.087	0.066	0.050	0.054
$f(a,\cdot) \times$ PRODLO	-0.00039	0.00084	0.00056	0.00063
$f(a,\cdot) \times$ PRODHI	0.0024	0.0012	0.00080	0.00075
$f(a,\cdot) \times$ EDUCATION	0.035	0.019	0.012	0.013
$f(a,\cdot) \times$ MATCH	0.0049	0.0082	0.0061	0.0086
$f(a,\cdot) \times$ JOB COMPLEXITY	-0.0071	.0097	0.0070	0.0063
MARRIED FEMALE	0.530	0.371	0.278	0.288
BOX-COX PARAMETER	0.182	0.155	0.144	0.198
DISCOUNT RATE (δ)	0.046	0.026	0.022	0.027
SPLIT (B)	143.398	14.378	0.939	0.075
CORRELATION: Quit. Ed. Errors	-0.595	0.255	0.166	0.178

FINAL Ln LIKELIHOOD = 359.753
(Full Model)

---

a   See Table 2 for definitions of $f(a, \cdot)$, PRODLO, PRODHI, and B.

TABLE 1: THE DATA

MARGINAL SUMMARY

	MEAN	STD. DEV.	MINIMUM	MAXIMUM
WHITE	0.7775378	0.4163498	0.0	1.000000
BLACK	0.1857451	0.3893215	0.0	1.000000
MALE	0.5572354	0.4972506	0.0	1.000000
AGE	24.24303	7.013603	17.51781	58.27122
EDUCATION	12.00831	1.105222	8.000000	19.00000
Ln(PREVIOUS WAGE)	0.6366167	0.1960310	0.2176941	2.096026
TENURE	0.538473	2.095973	0.01803831	13.27557
EXPERIENCE	3.448529	3.430775	0.1600455	23.53396
OUTPUT	111.8827	25.23188	16.55630	205.4802
PINS	22.36285	4.024556	11.00000	35.00000
SCREWS	22.73866	5.164266	8.000000	39.00000
TR. DUR. (WKS.)	7.611234	5.066129	1.000000	36.00000
MATCH	-0.0630041	0.8250197	-7.107641	3.724559
1977 COHORT	0.4600432	0.4989400	0.0	1.000000
QUIT	0.0669546	0.2502138	0.0	1.000000

CORRELATIONS

	WHITE	BLACK	MALE	AGE
WHITE	1.000000	-0.8929171	0.02504234	-0.00651530
BLACK	-0.8929171	1.000000	-0.03267315	-0.04412565
MALE	0.02504324	-0.03267315	1.000000	-0.2622108
AGE	-0.006515302	-0.04412565	-0.2622108	1.000000
EDUCATION	-0.06111908	0.05978535	0.09670897	-0.1554595
Ln(PREVIOUS WAGE)	0.08962641	-0.08501559	0.3542936	0.08447425
TENURE	-0.01241098	0.02062389	-0.02118682	0.3433512
EXPERIENCE	-0.01495738	0.005370273	-0.1007093	0.7156301
OUTPUT	0.06208870	-0.06659605	0.1351162	-0.05111940
PINS	0.2032888	-0.1895405	-0.1455995	-0.06227949
SCREWS	0.1178640	-0.1308295	0.2768285	-0.08536030
TR. DUR. (WKS.)	0.02612329	-0.01214396	-0.1763108	0.05486028
MATCH	-0.03273983	0.01268510	-0.006980062	-0.03709897
1977 COHORT	-0.07935104	0.1162912	0.1073870	-0.08020447
QUIT	-0.002154015	-0.01684474	0.03002183	-0.07405117

TABLE 1: CORRELATIONS (cont.)

	EDUCATION	Ln(PREVIOUS WAGE)	TENURE	EXPERIENCE
WHITE	-0.061	0.090	-0.012	-0.015
BLACK	0.060	-0.085	0.021	0.005
MALE	0.097	0.354	-0.021	-0.101
AGE	-0.155	0.084	0.343	0.716
EDUCATION	1.000	0.079	-0.136	-0.116
Ln(PREVIOUS WAGE)	0.079	1.000	0.164	0.149
TENURE	-0.136	0.164	1.000	0.565
EXPERIENCE	-0.116	0.149	0.565	1.000
OUTPUT	0.046	-0.005	-0.011	-0.032
PINS	0.066	-0.038	-0.053	-0.074
SCREWS	0.083	0.170	0.026	-0.002
TR. DUR. (WKS.)	-0.053	-0.113	-0.067	-0.011
MATCH	0.128	0.016	-0.048	0.002
1977 COHORT	0.017	-0.063	0.029	0.012
QUIT	-0.007	0.032	-0.023	-0.043

	OUTPUT	PINS	SCREWS	TR. DUR. (WKS.)
WHITE	0.062	0.203	0.118	0.026
BLACK	-0.067	0.190	-0.131	-0.012
MALE	0.135	-0.146	0.277	-0.176
AGE	-0.051	-0.062	-0.085	0.055
EDUCATION	0.046	0.066	0.083	-0.053
Ln(PREVIOUS WAGE)	-0.005	-0.038	0.170	-0.113
TENURE	-0.011	-0.053	0.026	-0.067
EXPERIENCE	-0.032	-0.074	-0.002	-0.011
OUTPUT	1.000	0.109	0.068	0.208
PINS	0.109	1.000	0.172	0.039
SCREWS	0.068	0.172	1.000	-0.042
TR. DUR. (WKS.)	0.208	0.039	-0.042	1.000
MATCH	0.064	0.003	0.003	0.042
1977 COHORT	0.137	0.106	0.109	0.059
QUIT	-0.086	-0.003	0.032	0.074

TABLE 1: CORRELATIONS (cont.)

	MATCH	1977 COHORT	QUIT
WHITE	-0.033	-0.080	-0.002
BLACK	0.013	0.116	-0.017
MALE	0.007	0.107	0.030
AGE	-0.037	-0.080	-0.074
EDUCATION	0.128	0.017	-0.007
Ln(PREVIOUS WAGE)	0.016	-0.063	0.032
TENURE	-0.048	0.029	-0.023
EXPERIENCE	0.002	0.012	-0.043
OUTPUT	0.064	0.137	-0.086
PINS	0.003	0.106	-0.003
SCREWS	0.003	0.109	0.032
TR. DUR. (WKS.)	0.042	0.059	0.074
MATCH	1.000	-0.050	-0.030
1977 COHORT	-0.049	1.000	-0.022
QUIT	-0.030	-0.022	1.000

TABLE 2 - RESULTS

	VALUE	GRADIENT SE	HESSIAN SE	WHITE SE
DEPVAR = Ln(PREV. PAY):				
CONSTANT	-1.4679892	0.5893389	0.5218230	0.5916407
TENURE	0.0127592	0.0075375	0.006726	0.0082664
EXPERIENCE	0.0030894	0.0075428	0.0055626	0.0048669
(AGE-16.)	0.0234616	0.0060018	0.0054978	0.0058606
(AGE-16.)**2	-0.0005684	0.0001686	0.0001554	0.0001849
EDUCATION (YRS.)	0.0559459	0.0490274	0.0433087	0.0498388
NONWHITE	-0.0551047	0.0410293	0.0305364	0.0286057
MALE	0.2224639	0.0321326	0.0266867	0.0253156
PINS	-0.0007896	0.0036820	0.0030543	0.0027956
SCREWS	0.0029098	0.0026118	0.0023774	0.0025354
MARRIED	0.0372976	0.0269069	0.0238405	0.0242925
1977 COHORT	-0.0679137	0.0276155	0.0234923	0.0228845
DEPVAR=Ln(OUTPUT):				
CONSTANT	4.3139129	0.6620291	0.4893330	0.4746207
EXPERIENCE	-0.0001923	0.0057970	0.0047656	0.0047144
AGE	-0.0002275	0.0034783	0.0026547	0.0024060
Ln(EDUCATION)	0.0064917	0.0542153	0.0401697	0.0383926
NONWHITE	-0.0258299	0.0373514	0.0288854	0.0257058
MALE	0.0932603	0.0315719	0.0261182	0.0254149
PINS	0.0065033	0.0035376	0.0029339	0.0028392
SCREWS	-0.0005086	0.0028145	0.0022711	0.0021387
MATCH	0.0168088	0.0181104	0.0140501	0.0142590
JOB COMPLEXITY	0.0563695	0.0199372	0.0152505	0.0151799
1977 COHORT	0.0428478	0.0298555	0.0232790	0.0222756
DEPVAR=Ln(EDUCATION):				
NON-BLACK MALE	2.4782963	0.0111142	0.0095167	0.0104603
NON-BLACK FEMALE	2.5150405	0.0138290	0.0101989	0.0109479
BLACK MALE	2.4834337	0.0510206	0.0278454	0.0202355
BLACK FEMALE	2.4979623	0.0259124	0.0210335	0.0227843
NON-BLACK MALE × AGE	0.0013340	0.0009636	0.0011058	0.0016833
NON-BLACK FEMALE × AGE	-0.0048816	0.0006214	0.0007660	0.0013499
BLACK MALE × AGE	0.0016551	0.0063807	0.0044623	0.0042272
BLACK FEMALE × AGE	-0.0002344	0.0016029	0.0017927	0.0027359
ASIAN	0.0783867	0.0366227	0.0324142	0.0355433
HISPANIC	0.0045484	0.0243293	0.0245449	0.0272952

DEPVAR=QUIT:

CONSTANT	-8.3292984	2.8349521	1.6836925	1.7694591
$f(a, \cdot) \times$ Ln(ADJ. PREVIOUS WAGE)[a]	0.0708688	0.0676113	0.0536747	0.0517533
$f(a, \cdot) \times$ PRODLO[b]	-0.0004947	0.0010442	0.0006280	0.0004769
$f(a, \cdot) \times$ PRODHI	0.0023541	0.0013478	0.0009081	0.0008293
$f(a, \cdot) \times$ EDUCATION	0.0525292	0.0243542	0.0197303	0.0240889
$f(a, \cdot) \times$ MATCH	0.0012942	0.0098327	0.0052634	0.0055623
$f(a, \cdot) \times$ JOB COMPLEXITY	-0.0049813	0.0101997	0.0067884	0.0060454
MARRIED FEMALE	0.3413957	0.2143069	0.1717866	0.1662031
DISCOUNT RATE ($\cdot$)	0.0733651	0.0263590	0.0217058	0.0252488
SPLIT (B)	142.4914782	16.1064599	13.2154879	12.3572205

DISTRIBUTION VARIANCES AND
CORRELATIONS OF ERROR TERMS

	VALUE	GRADIENT SE	HESSIAN SE	WHITE SE
CORR(Ln(WAGE),QUIT)	0.0059954	0.2644620	0.2324430	0.2542600
CORR(Ln(OUTPUT),QUIT)	-0.3582333	0.3465599	0.2011844	0.1665908
CORR(Ln(EDUCATION),QUIT)	-0.7579667	0.1777674	0.1335131	0.1492422
VARIANCE(Ln(PP) ERROR)	0.0628291	0.0054896	0.0058066	0.0080895
CORR(Ln(W),Ln(OUTPUT))	-0.0739408	0.0816091	0.0558964	0.0449459
CORR(Ln(WAGE),Ln(EDUCATION))	-0.1810289	0.1919123	0.1798617	0.2129612
VARIANCE(Ln(OUTPUT) ERROR)	0.0598328	0.0026746	0.0039342	0.0092396
CORR(Ln(OUTPUT),Ln(EDUCATION))	-0.0046549	0.2397400	0.1786869	0.1785935
VARIANCE(Ln(EDUCATION) ERROR)	0.0079264	0.0004022	0.0005218	0.0008537

FINAL Ln (LIKELIHOOD) = 358.797412

---

a   The term $f(a, \cdot)$, which depends on age and the discount rate, is derived from equations (2)-(3).

b   The variables PRODLO and PRODHI refer to low and high productivity measures respectively. With B
as a parameter denoting an output split:

$$PRODLO \equiv \begin{cases} [B - \text{Output}], & \text{for Output} < B \\ 0 & \text{otherwise} \end{cases}$$

$$PRODHI \equiv \begin{cases} [\text{Output} - B], & \text{for Output} > B \\ 0 & \text{otherwise} \end{cases}.$$

REFERENCES

Arnold, Hugh and Daniel Feldman, "A Multivariate Analysis of the Determinants of Job Turnover," *Journal of Applied Psychology,* Vol. 67, No. 3, (June 1982) pp. 350-360.

Bartel, Ann, "Wages, Nonwage Job Characteristics and Labor Mobility," *Industrial and Labor Relations Review,* Vol. 35, No. 4 (July 1982) pp. 578-589.

Bloch, Farrell, "Labor Turnover in U. S. Manufacturing Industries," *Journal of Human Resources,* Vol. 14, No. 2 (Spring 1979) pp. 236-245.

Flinn, Christopher, "Wages and Job Mobility of Young Workers," unpublished mimeo 1983.

Garen. John, "The Returns of Schooling, A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica,* Vol. 52, No. 5 (September 1984) pp. 1199-1218.

Hausman, J. A., "Specification Tests in Econometrics," *Econometrica,* Vol. 46, No. 6 (November 1978) pp. 1251-1272.

Heckman, James, "The Common Structure of Statistical Models of Truncation. Sample Selection, and Limited Dependent Variables and a Simple Estimator for such Models," Annals of Economic and Social Measurement, Vol. 5, (1976) pp. 475-492.

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica,* Vol. 47, (1979) pp. 73-92.

Knapp, Martin; Kostas Harissis and Spyros Missiakoulis, "Investigating Labor Turnover and Wastage Using the Logit Technique," *Journal of Occupational Psychology,* Vol. 55, (1982) pp. 129-138.

Mellow, Wesley, "Employer Size and Wages" *Review of Economics and Statistics,* Vol. 64, No. 3 (August 1982) pp. 495-501.

Mobley, W. H., Griffeth, H. H. and Meglino, B. M., "Review and Conceptual Analysis of the Employee Turnover Process", *Psychological Bulletin,* (1979) Vol. 86, pp. 493-622.

Pencavel, John, "Wages, Specific Training and Labor Turnover in U. S. Manufacturing Industries," *International Economic Review,* Vol. 13, No. 1, (1972) pp. 53-64.

Polachek, Solomon, "Occupational Self-Selection: A Human Capital Approach to Sex Differences" *Review of Economics and Statistics,* Vol. 63, No. 1, (February 1981) pp. 60-69.

Steers, R. and S. Rhodes, "Major Influences on Employee Attendance: A Process Model" *Journal of Applied Psychology,* (1978) Vol. 63, No. 4, pp. 391-407.

Viscusi, Kip "Sex Differences in Worker Quitting" *Review of Economics and Statistics,* Vol. 62, No. 3 (August 1980) pp. 388-398.

Weiss, Andrew, "Determinants of Quit Behavior for Semi-skilled Production Workers", *Journal of Labor Economics,* Vol. 2, No. 3. (July, 1984) pp. 371-387.

Weiss, Andrew "Testing the Sorting Model of Education," Columbia University discussion paper No. 245, (April 1984).

Weiss, Andrew "High School Graduation, Performance and Earnings": NBER Working Paper No. 1595, (April 1985a).

Weiss, Andrew "Wages and Absenteeism'" *Economics Letters*, No. 19, (1985b) pp. 277-279.

White, Halbert "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, Vol. 50 (1982) pp. 1-25.

White, Halbert "Corrigendum," *Econometrica*, Vol. 51, No. 2 (March 1983) pp. 513.