The Importance of Obtaining a High-Paying Job

Devereux, Paul J.

University College Dublin

January 2002

THE IMPORTANCE OF OBTAINING A HIGH PAYING JOB*

Paul J. Devereux
Department of Economics
University of California, Los Angeles
devereux@econ.ucla.edu

January, 2002

Abstract

Given the high level of job mobility in the United States, one might think that obtaining a low-paying job would have only temporary consequences. However, using longitudinal data, I find that state dependence in wages is large and persistent. If two comparable individuals start jobs that pay a different wage, about 60% of the wage differential is still present four years later. Moreover, about 50% of the wage differential is still present for workers who have switched employers during that period. The results indicate that the jobs acquired by individuals have long-term effects on their future careers. I also examine the mechanisms that lead to state dependence. In a stigma model, prospective employers use wages as a signal of ability. Thus, getting a poor job can lead the market to believe that an individual has low ability. In the learning-by-doing model, workers who get high-paying jobs also attain greater opportunities to acquire human capital. The evidence suggests that both stigma and learning by doing models contribute to state dependence in wages.

* I thank Joseph Altonji, Guido Imbens, Dale Mortensen, Chris Taber, Sandy Black, Enrico Moretti, Phillip Leslie, Dan Ackerberg and Kanika Kapur for helpful conversations. I also thank seminar participants at Northwestern, Rochester, UCLA, Washington, Maryland, Yale, and at the Federal Reserve Bank of Atlanta. This research was supported by an Alfred P. Sloan Doctoral Dissertation Fellowship. A previous version of this paper circulated under the title "State Dependence in Wages: Stigma, Human Capital, or Search?".

1. INTRODUCTION

Recent contributions in information theory have suggested that taking employment at an unskilled job serves as a bad signal for firms hiring workers to skilled jobs.[1] These models imply that there is state dependence in job quality in that temporary shocks that affect the types of jobs workers obtain have large and persistent effects on subsequent wages. However, given the high level of job mobility in the United States, it is not clear that job assignment should have long-term effects. In this paper, I examine the importance of state dependence in the types of jobs workers hold over their careers. One can think of the central issues in terms of the following thought experiment. Imagine a group of people who are qualified to do a range of different jobs and hence earn a range of different wages. These people are randomly assigned a job within the range of jobs that they are qualified to do. If state dependence in wages is important, the initial placement the worker receives has large and persistent effects on subsequent job quality and wages. In this paper, I show that state dependence in job quality is an important labor market phenomenon. Given that there is evidence of significant persistence, I then focus on some potential explanations for it.

The principle problem in estimating the degree of state dependence is individual heterogeneity: Individuals who get low-paying jobs are typically less skilled. However, the standard method of dealing with unobserved heterogeneity, fixed effects, is inconsistent when there is a lagged dependent variable. Therefore, to estimate the importance of state dependence, I use instrumental variable techniques. I exploit the fact that the assignment of workers to jobs changes over the business cycle. Therefore, I use the state unemployment rate at the time of hire as an instrument for the wage and characteristics of the job obtained. A series of specification checks indicate that the assumption that the starting state unemployment rate is uncorrelated with unobserved ability is a valid one.

The second major question this paper answers is why shocks to job quality should have persistent effects. I consider several possible mechanisms. The first is simply that jobs persist and so long as an individual stays with the same employer, their wage will not change very much. The second mechanism is a simple search explanation in which new offers arrive at intervals and so it takes time for individuals who start with a low wage to catch up to persons who start with a high wage.[2] Empirically, I show that it is unlikely that these factors by themselves explain all state

[1] See Ma and Weiss (1993) and McCormick (1990).
[2] Mortensen (1986) provides a survey of the search literature.

dependence in wages. Thus, I consider stigma and learning by doing explanations for state dependence in job quality.

The stigma model implies that the job held affects the market's perception of a worker's abilities. Theoretical papers by McCormick (1990) and Ma and Weiss (1993) suggest that taking a bad job may be a negative signal to employers. Because the market does not perfectly observe ability, it uses job quality as a signal of ability. Thus, exogenously attaining a job that is poor relative to ability stigmatizes workers and makes it more difficult for them to get good jobs in future. In this paper, I utilize methods pioneered by Farber and Gibbons (1996) and Altonji and Pierret (2001) to test implications of the stigma model. As such, I extend the recent empirical literature that studies the effects of employer learning on wages to the case where employers draw inferences about ability from wages.[3]

In contrast, the learning by doing explanation is that one's skills are changed by the job one gets: Good jobs have greater potential for human capital accumulation, and therefore achieving a high wage can be positively correlated with opportunities to acquire skills. In this model skill accumulation is a byproduct of working on the job. Even if workers must pay for learning opportunities, it is possible for a worker to obtain a job that involves both a higher wage and greater learning opportunities. For example, if a worker finds a good match in which he is very productive, he may receive a higher wage than if he had found a poor match even after paying for the greater learning by doing opportunities available. Thus, if a worker exogenously gets a job that is better than his ability would suggest, he may increase his ability through learning by doing. Therefore, the job currently held affects his future ability and, hence, the type of job he holds in the future.

The extent of state dependence in wages is important for many reasons. State dependence implies that there is a high payoff to labor market search even when job tenures are generally short, and hence is an explanation for voluntary unemployment. The strategy of taking a poor job and searching on the job may be dominated by continuing to search from unemployment, even if search costs from unemployment and employment are similar. Also, the question is relevant to assessing the importance of policies designed to facilitate the matching of workers to jobs that best exploit their skills. The causes of state dependence are also potentially important to distinguish. If workers are stigmatized by taking poor jobs, aggregate fluctuations that affect assignment in the labor market can reduce future economy productivity by distorting information flows about worker ability. If, instead, state dependence results from greater human capital accumulation on good jobs, then this implies that human capital models should place more emphasis on the constraints that workers face in deciding whether to invest. It would also indicate the importance of retraining for workers who have been displaced to less-skilled jobs.

Before proceeding further, it is useful to discuss the relationship between this study and the previous literature. There is little research that studies the importance of job assignment to future wages or that empirically investigates the mechanisms that cause such persistence. There is, however, a related literature on wage dynamics and the covariance structure of earnings.[4] The emphasis of these papers differs from that of the current paper in that they do not test for causal relationships between current and future wages. Rather, the models in these papers assume that there is no state dependence in earnings but that earnings residuals are correlated over time. Thus, a common interpretation is that the wage residual measures unobserved ability, and negative first order autocovariances in wage changes reflect investment in human capital and subsequent payoff. There is a fundamental difference between state dependence and serially correlated errors. With state dependence the job one gets causally affects the jobs held in future, with serially correlated errors, there is just some correlation over time in the wages earned by workers. Note that the earnings dynamics literature typically does not model jobs and job changes in any way. Thus the results from this literature do not in any way imply that there is a persistent benefit to getting a high-paying job.

Another related paper is that by Beaudry and DiNardo (1991). They show that the national unemployment rate when a worker starts a job has a negative effect on wages on that job even after conditioning on the current unemployment rate. This paper expands on their work by quantifying the effect of getting a high-paying job on future wages without conditioning on the worker staying with the same employer: Indeed, I show that there is state dependence in wages even when workers switch jobs. I also add to the literature by exploring stigma and human capital accumulation mechanisms for state dependence.

The structure of this paper is as follows: In section 2, I describe the data set. In section 3, I discuss the econometric difficulties in estimating dynamic models with individual heterogeneity, outline the techniques that I use, and present the empirical results for the state dependence specifications. In section 4, I describe a series of specification checks for the state dependence

---

[3] Farber and Gibbons (1996) and Altonji and Pierret (2001) study the signaling value of education in the context of a learning model in which the market learns about worker productivity by observing output realizations.

[4] See Kearl (1988), Hause (1980), Parent (1995), Abowd and Card (1989), Topel and Ward (1992), Farber and Gibbons (1996), MaCurdy (1982), Holtz-Eakin, Newey, and Rosen (1988), and Geweke and Keane (1997) amongst others.

estimates. In section 5, I examine the possible state dependence mechanisms including search, stigma, and learning by doing. Section 6 summarizes the results and concludes.

2. DATA

The data used throughout this paper come from the 1971 to 1992 survey years of the Panel Study of Income Dynamics (PSID). This data set was chosen because it extends over a long time period and is representative of the working age population of the United States. The PSID is composed of both a random sample and a poverty subsample. I use observations from both samples to increase the precision of the estimates; the estimates are similar when just the random sample is included. I restrict the sample to men between the ages of 21 and 55 who are not self-employed and are not retired. I exclude observations from Hawaii and Alaska. In the analysis to estimate the importance of state dependence, I only use post-1976 observations. However, data before 1976 are used when previous values of variables are required.

There are two potential wage measures in the PSID: the reported hourly wage rate and annual average hourly earnings. I have chosen to use the reported wage as it is specific to the current job. For hourly workers this variable is the reported hourly wage. For workers paid weekly, monthly, or annually, it is the amount paid divided by a fixed number of hours. For example, for weekly workers, the wage rate is created by dividing weekly earnings by forty hours. Since the hourly wage is topcoded at $9.98 in 1976 and 1977, many wage observations in those years are topcoded. In these cases average hourly earnings for the year have been used to predict the wage. Wages are deflated by the GDP consumption deflator.

The analysis in this paper requires the identification of job moves. Two tenure questions are asked in the data: 'How long have you been in your present position?' and, 'how long have you worked for your current employer?'. These data allow one to determine when employer changes, and position changes within firms, occur. I partition the data into spells with employers using the method recommended by Brown and Light (1992). An individual is assumed to have started a spell with a new employer when tenure with the employer is less than the elapsed time since the survey date. Before 1976 the tenure data in the PSID is bracketed. If reported tenure in these years is less than one year then I assume that the worker has started a new job. If tenure is greater than one year, I assume that the job is the same as the job held in the previous year. Table 1 contains means of selected variables.

One strength of the PSID data is that respondents who report less than one year of tenure on their current job are asked about what happened to their previous job. The responses are coded as being a quit, a displacement because of plant closing, or as a firing/layoff. As we will see, it turns out to be useful to distinguish between voluntary and involuntary transitions in the analysis.

3. MEASURING STATE DEPENDENCE IN WAGES

The Estimation Equation

In this section I examine whether the type of job a worker obtains has long-term career implications. I take a sample of job starters and analyze how the job they hold up to 5 years later depends on the type of job they start. I choose a time period of five years as this is a long enough period to be considered long-term but is short enough for me to have a reasonable number of observations on workers present in both periods. This question is central to understanding the importance of assignment to future wages in the labor market. If the assignment a worker obtains still has large effects on wages several years later, it implies that temporary shocks that affect assignment have large effects on the careers of workers. I use the wage on the job as a proxy for the quality of the job. There are other job characteristics that are important to workers but wages are generally assumed to be the most important.

Suppose we observe a group of people starting new jobs at time t. We are interested in the extent to which the wage they attain at time t affects the wage they earn some time later, say, at time τ. To answer this descriptive question, it is quite natural to write a specification such as equation (1)

$$
\begin{aligned}
w_{i\tau} &= \beta_1 w_{it} + \beta_2^{\prime} x_{it} + \beta_3 u_{\tau} + \beta_4 u_{\tau-1} + v_{i\tau} \qquad (1) \\
v_{i\tau} &= f_i + \eta_{i\tau}
\end{aligned}
$$

Here, w_{it} is the log of the wage of person i at time t, x_{it} is a vector of control variables, and u_{τ} is the state unemployment rate at τ. The vector of control variables include personal characteristics at t, the log of the wage at t, a full set of state indicators, and a quadratic time trend. The individual characteristics I include are labor actual market experience, experience squared, experience cubed, and indicators for whether married, whether has a college degree, whether has a high school diploma, and an indicator for whether white. I report OLS estimates to provide a comparison point for the later estimates. The results are in table 2. The OLS estimates show a decline in the coefficient on w_{it} from about 0.7 at t+1 to about 0.57 at t+4 and t+5.

Estimation Issues

The error ($v_{i\tau}$) is assumed to be composed of an individual effect (f_i), and a, possibly serially correlated, idiosyncratic component $\eta_{i\tau}$. The x_{it} and u_τ variables are assumed to be strictly exogenous given the unobservable individual effect, f_i. There are a number of econometric problems that cause OLS estimates of equation (1) to be biased. The principal problem is the correlation of the wage with the individual effect because people with greater ability will tend to have better jobs. Because of this unobserved heterogeneity, there will be a spurious positive correlation between $w_{i\tau}$ and w_{it} that arises because $Cov(w_{it}, f_i) > 0$. This will tend to bias the estimate of β_1 upwards.

The presence of individual effects that are potentially correlated with the regressors is a serious problem in a dynamic model such as model (1). The "fixed effect" response of OLS on deviations from individual means is inconsistent in dynamic models because taking deviations from individual means induces a correlation of order 1/T between the lagged dependent variable and the error (Nickell, 1981). The most common response to this problem is to first difference the equation to remove the effects, and then to estimate by instrumental variables, using w_{it-1} and further lags of the dependent variable as instruments for Δw_{it} (Anderson and Hsiao, 1982). In the absence of measurement error and serially correlated errors, this method produces consistent but inefficient estimates. Recent Monte Carlo results (Arellano and Bover, 1995) suggest that small sample biases can be very substantial when the variance of the individual effect and the coefficient on the lagged dependent variable are large. In the presence of serially uncorrelated measurement error in $w_{i,t-1}$, only $w_{i,t-2}$ and further lags of the wage are valid instruments. In this case the following moment conditions hold:

$$E(w_{i,t-2}\Delta v_{i\tau}) = 0 \text{ and } E(w_{i,t-2}\Delta w_{it}) < 0$$

The explanatory power of the levels in wages in the presence of measurement error is small, so with measurement error one is likely to suffer more serious biases than these Monte Carlo results suggest. For this reason, I do not use this method in the estimation.

Using Lagged Wage Changes as Instruments

Arellano and Bover (1995) suggest applying predetermined, but not necessarily exogenous, variables in the level equation as first differenced instruments. If the model is stationary and the correlation between the predetermined variables and the individual effect is time invariant, these variables are valid instruments.[5] In particular, the first difference of the lagged dependent variable may be used as an instrument (that is, Δw_{it} may be used as an instrument for w_{it}). The following moment conditions apply: $E(v_{i\tau}\Delta w_{it}) = 0$ and $E(w_{it}\Delta w_{it}) > 0$. This is the first instrument I use in this analysis. The results using this instrument are reported in the second row of table 2. As can be seen in the second column, this instrument has a strong first stage. The 2SLS result is that the coefficient on w_{it} is 0.187 (0.045) after four years. This estimate is much lower than the corresponding OLS estimate. However the instrument used is not robust to the presence of measurement error in the wage because the measurement error in the instrument is correlated with the measurement error in w_{it}. Thus the 2SLS estimates are likely to be downwardly biased.[6]

With serially uncorrelated measurement error, Δw_{it} can be replaced as an instrument by Δw_{it-1}. In row 3 of table 4, the instrument for w_{it} is $w_{it-1} – w_{it-2}$, while in row 3 the instrument is the log of average hourly earnings at t (e_{it}) minus the log of average hourly earnings at t-1. The assumption with this latter instrument is that measurement error in average hourly earnings is uncorrelated with measurement error in the reported wage.[7] Because of measurement error, the relationship between the instrument and the predetermined variable is much weaker in rows 3 and 4 than in row 1. However, the instrument is still highly correlated with w_{it}. Using these instruments changes the estimate of state dependence radically. The coefficient at t+4 rises to 0.616 (0.176) when $w_{i,t-1} – w_{i,t-2}$ is used as an instrument, and to 0.646 (0.137) when $e_{it} – e_{i,t-1}$ is used as the instrument. The change in the coefficient on w_{it} indicates the importance of taking account of measurement error in the wage. In the t+1 regression, the coefficient on w_{it} is greater than one when $e_{it} – e_{i,t-1}$ is used as the instrument. This probably occurs because average hourly earnings at t may include earnings from many jobs held in that year. Thus, if an individual changes jobs during year t, e_{it} is a function of the wage of the job held at t+1 and the instrument is correlated with the error term. This should not be a serious problem for the regressions subsequent to t+1.

[5] This stationarity assumption implies that there is zero covariance between the individual effect and wage changes. This assumption is consistent with the findings of Abowd and Card (1989) and MacCurdy (1992) that use the PSID data. However, Baker (1997) provides evidence that wage growth is faster for more highly paid workers.
[6] Validation studies have found measurment error in earnings in the PSID to be quite severe (Bound, Brown, Duncan, and Rodgers, 1994). To my knowledge, the wage measure used in this paper has not been validated.
[7] Since the information on reported wages and the information on earnings for any particular year are collected in different years, measurement error may not be highly correlated in the two measures. Average hourly earnings is computed over the calender year and so, unlike the reported wage, it is not a point in time measure.

Using the State Unemployment Rate

The wage instruments are not robust to certain types of serial correlation in the error term. If, for example, a worker becomes more productive and so gets a high w_{it} relative to his previous wages, then the error term will be positively correlated with Δw_{it}. This is likely to bias the estimates upwards. A similar problem arises if there are changes in worker tastes as he starts a new job. If the changes in tastes are persistent, then the error is positively correlated with w_{it}. Another possibility is that a worker may invest in skills on one job (and accept a lower wage to finance the general training) in order to learn skills required for a better one. In this case the error will be negatively correlated with Δw_{it}. Estimates are biased upwards if there is positive serial correlation in the errors; they are biased downwards if the serial correlation is negative. Therefore, positive coefficients on the lagged wage using the wage instruments may be because of serial correlation in the errors rather than state dependence. [8]

Therefore, I also use the state unemployment rate (u_t) as an instrument. The intuition is that the state unemployment rate at t affects the type of job a searching worker attains at t. One can think of the state unemployment rate as exogenously affecting the types of jobs that searching workers obtain.[9] Therefore, the unemployment rate when the job held at t started is correlated with the quality of job started at t. If the starting unemployment rate is uncorrelated with ability, then it will only affect $w_{i\tau}$ through its effects on w_{it} and, hence, it is a valid instrument for w_{it} provided that one controls for u_τ and $u_{\tau-1}$. [10] Two stage least squares estimation with the starting state unemployment rate instrument is robust to measurement error in the wage because measurement error in the starting unemployment rate is unlikely to be correlated with measurement error in the wage. The instrument is valid if it is not correlated with the error term ($v_{i\tau}$).

In the fifth row of table 2, I include the estimation results when the starting state unemployment rate is used as the instrument for w_{it}. The maintained assumption in these rows is

---

[8] Even with significant serial correlation, biases need not be enormous: For example, suppose half the variance of wage changes occurs due to random shocks, and the other half is the result of a serially correlated AR(1) process. Consider a case where there is no true state dependence and we use the Δw_{it-1} instrument. If the autoregressive coefficient is 0.8, then the estimated coefficient on w_{it} would be 0.32 at t+1 falling to 0.13 at t+5. If the autoregressive coefficient is 0.5, the estimated coefficient on w_{it} would be 0.13 at t+1 and 0.01 at t+5.

[9] As suggested by Beaudry and DiNardo (1991), cyclical conditions may also affect the wages of job starters because they affect the implicit contracts that agreed upon between workers and firms. According to this interpretation, cyclical conditions when the match start may affect wages but not affect the types of jobs that workers are doing.

that the state unemployment rate is uncorrelated with unobserved ability or tastes. The first stage shows that this instrument is strongly negatively correlated with w_{it}. The estimated levels of state dependence after 4 and 5 years are similar to when wage differences are used as instruments. The conclusion of strong state dependence remains. The estimates that use the state unemployment rate instrument are my preferred results because these are robust to measurement error in wages, serial correlation in the errors and to breaches of the assumption that the wage process is stationary.[11] One should not take the estimate for t+1 seriously when the state unemployment rate instrument is used because the estimating equation includes u_t and this is extremely highly correlated with the state unemployment rate when the job at t started.

There is a fundamental problem with the starting unemployment rate instrument if stigma is the reason for state dependence. If the market has any signal about the starting unemployment rate of the worker, they will put some weight on this information in determining the workers ability. Thus, in this scenario, the starting unemployment rate belongs in the regression, and if it is used as an instrument, it is correlated with the error. The stigma story implies that this correlation will be positive. Therefore, the coefficient on w_{it} will be biased downwards.[12]

Is there state dependence in Job types as well as wages?

Thus far, wages have been used, at least in part, to proxy for job quality. To verify that there also is state dependence in the types of jobs people hold, I construct a variable called the predicted wage. The predicted wage is the wage predicted from job characteristics. I create this variable by regressing the log wage on indicators for two-digit industry, two-digit occupation, union coverage, salaried, and government for the full sample that includes job starters and job stayers. The predicted wage is the predicted value from this regression. I use this variable because it complements the information about persistence in wages by indicating the extent to which job characteristics such as industry and occupation are persistent. The results using this variable are in table 3. On the whole, the estimates are quite similar but less precise than the results for wages. One exception is that the estimate using the state unemployment rate instrument for t+5 shows no evidence that job characteristics persist for 5 years. The other estimates in the table, however, do suggest substantial persistence over 4 and 5 years.

---

[10] These controls ensure that the starting unemployment rate does not affect wages due to correlations between the starting unemployment rate and the unemployment rate at τ.

[11] The reported standard errors take account of correlations in the error resulting from multiple observations on individuals.

4. SPECIFICATION CHECKS FOR THE STARTING UNEMPLOYMENT RATE INSTRUMENT

Solon, Barskey, and Parker (1994) conclude that the average ability of the workforce is higher in recessions. If it is the case that the average ability of job starters is also higher when the state unemployment rate is high, then this would bias the estimates of state dependence using the state unemployment instrument downwards. Below, I carry out some checks that suggest that the starting unemployment rate is not correlated with ability to any significant extent.

First, I check whether the starting unemployment rate is correlated with observable characteristics of workers on the basis that if it is correlated with observables it is probably correlated with unobservables. I regress the starting state unemployment rate on the state indicators, the quadratic time trend, and on individual characteristics. The coefficients on the individual characteristics are reported in the table below:

Married	0.032 (0.046)
College Degree	0.056 (0.057)
High School diploma	-0.039 (0.047)
Experience	-0.019 (0.025)
Experience Squared	0.0003 (0.0017)
Experience Cubed	0.00001 (0.00003)
White	0.058 (0.047)

As can be seen, none of the variables in this regression has any predictive power for the starting unemployment rate. An F test confirms the visual impression that one cannot reject that all these coefficients are zero at conventional significance levels. Thus, the starting unemployment rate appears to be random with respect to observed characteristics of individuals.

If the 2SLS estimates in table 2 are biased upwards because the starting unemployment rate is negatively correlated with unobserved ability, then one should find positive coefficients on

[12] It may be reasonable to assume that employers do not know the cyclical conditions that existed when a worker was hired on a previous job. To obtain this information, an employer must know exactly when the employee was hired, and what the unemployment rate was at that time.

w_{it} when the dependent variable is a previous value of the wage and one uses the starting unemployment rate as the instrument. The 2SLS estimate from the regression of w_{it-2} on w_{it} is -0.167 (0.287).[13] The equivalent estimate for the predicted wage is 0.022 (0.288). Failure to find a significantly positive coefficient in this regression suggests that the starting unemployment rate is not negatively correlated with ability so the estimates in tables 2 and 3 are not biased upwards. As is appropriate in this quasi-experimental setting, conditional on the observables, wages prior to the job start are unrelated to the starting unemployment rate. Then, wages when the job starts and subsequently are negatively related to the starting unemployment rate.

<u>Additional Specification Checks:</u>

Table 4 contains estimates using the starting unemployment rate instrument from several different specifications of the dynamic wage model. As can be seen in the table, the estimates are quite robust to specification. I outline the rationale behind these checks in detail below.

Adding Interactions of the personal characteristics with the quadratic time trend:

The period under study was one of rising wage inequality. If there is some spurious correlation between changes in relative wages and the state unemployment rate, these relative wage changes could lead to bias. I allow the education, experience, and race premiums to vary over this time period by interacting these variables with the quadratic time trend. This has little effect on the estimates

Adding Year Indicators:

I replace the quadratic time trend with a full set of year indicators in case there are year specific-factors that affect wages and are correlated with the starting state unemployment rate. One practical effect of adding the year dummies is to condition out variation in the state unemployment rate that is correlated with the national business cycle. The first stage remains strong but the 2SLS estimates are less precisely estimated than in table 2. The point estimates of state dependence are higher than in table 2.

Adding Controls for Voluntary or Involuntary Separation:

Workers are more likely to experience involuntary separations in recessions. If laidoff workers are of lower quality, this could induce a negative relationship between the starting

[13] I don't do the equivalent regression for w_{it-1} because the starting unemployment rate is extremely highly correlated with the state unemployment rate at t-1.

unemployment rate and ability. As a check on this possibility, I add controls for whether workers started their previous job because of a quit, or started their previous job because they were involuntarily separated. With these extra controls, I find estimates of state dependence that are similar in magnitude to the estimates in table 2. Thus, it seems that the greater likelihood that job starters in recession are separated involuntarily is not exerting a serious bias on the state dependence results.[14]

Adding Controls for the Average unemployment rate of the individual and the average starting unemployment rate of the individual:

Solon, Barskey, and Parker (1994) report that the skill composition of the employed changes over the business cycle. I take the correlated random effects approach of specifying the individual effect as a function of the mean value of the state unemployment rate for the individual. To do this, I calculate the mean value of the state unemployment rate for each period in which the person is employed at the interview date. I also calculate the mean starting state unemployment rate for each individual. Then, these two variables are added as extra controls in the state dependence specifications. Once again, the estimates do not change very much.

Excluding the Poverty Sample:

One further objection to the results in table 2 is that the sample includes both the random sample and the poverty sample components of the PSID. Therefore, I present the estimates when the poverty sample is excluded. Once again, the results are quite similar to the results in table 2. The point estimates on the coefficient on w_{it} are slightly higher then with the full sample but, given the precision of the estimates, it is difficult to see how the amount of state dependence differs between the random sample and the full sample.

Using the same individuals in each year:

One feature to note about table 2 is that there are different numbers of cases in each year. In row 6 of table 4, I restrict the sample to cases where the relevant variables are present in every year to ensure that the estimated profile is not contaminated by having a different sample in each year. Reassuringly, the profile when the same observations are used in each year is quite similar to the profile in table 2.[15]

[14] I have also tried including lagged tenure and interactions of lagged tenure with the voluntary and involuntary termination variables. The estimates remain very similar.

[15] Similarly, I have verified that differences in the estimates across the wage and starting unemployment rate instruments in table 2 are not the result of differing samples.

Weighting the Sample to Correct for Attrition and Non-Employment:

Many wage observations are missing because of attrition and non-employment at the survey date. Using a selection on observables approach, I weight each observation by the inverse of the predicted probability of being present in the sample. I report estimates in table 4 using weights that correct for attrition only, and weights that control for both non-employment and attrition. Details about the attrition problem and the methodology used to address it are in appendix 1. In the first set of results the weights are the inverse of the estimated probability of not attriting from the PSID between t and τ, in the second set of results the weights are the inverse of the estimated probability of having complete data at τ. The weighted 2SLS estimates are similar to the estimates in table 2 but the point estimates are generally lower. This reflects the fact that individuals who have characteristics that make them more likely to attrit or be nonemployed have less persistent wages.

Replacing missing wages with average hourly earnings:

There are many cases where the individual remains a PSID respondent but wage data is missing (see appendix 1 for details). In the final row of table 4, I report results where average hourly earnings is used to proxy for missing wages. For example, at t+4, of the 1103 cases where the wage is missing, there are valid observations on average hourly earnings at t+4 for 853 cases. In this manner, it is possible to check whether missing wages that do not result from attrition are biasing the results. The exercise indicates that it is unlikely that selection due to missing wages has a major effect on the estimates of state dependence in table 2.

<u>Conclusions on the extent of State Dependence in Wages:</u>

Since the results that are robust to measurement error all have reasonably large levels of state dependence, my conclusion is that getting a bad job has large and persistent effects on the subsequent careers of workers. The estimates with the starting state unemployment rate instrument imply that if two identical workers start different jobs, approximately 60% of the wage differential remains four years later. This is a substantial amount of state dependence. Since only one third of the workers are in the same job at t+4 that they started at t, one might have thought that the job started would not be so important. However, the results imply that cyclical shocks that affect the assignments workers receive have major long-term implications in labor markets. Given the finding of state dependence in wages, I now examine why this substantial state dependence might arise.

5. STATE DEPENENCE MECHANISMS

Is State Dependence Purely Because Jobs Persist?

One explanation for state dependence in wages is that, if a worker gets a high wage job, the wage of that worker remains high over time because he remains on the same job. Furthermore, there may be implicit contracts or nominal rigidities that cause wages within jobs to be particularly persistent. Thus, before moving on to discuss other explanations for state dependence, it is important to determine that state dependence in wages does not purely result from the fact that jobs persist. To examine this issue, I estimate the regressions on a sample of individuals who have changed employers at least once between t and τ. Using the state unemployment rate instrument, I find that the coefficient on w_{it} is 0.566 (0.210) when the dependent variable is w_{it+4} and the sample is restricted to individuals who have switched employers between t and t+4. The analogous coefficient on w_{it} in the t+5 regression is 0.545 (0.253). The estimates from these regressions are also around 0.5 when the wage instruments that are robust to measurement error are used. Thus, the effect of getting a high wage job at t persists even if workers subsequently switch jobs.

Can a Search Model Explain all State Dependence in Wages?

Even without human capital accumulation or stigma, the quality of job obtained may influence future job quality by affecting the reservation wage required to switch to a new job. In this search explanation wage offers arrive periodically. Thus, it takes time for individuals who start with a low wage to catch up to persons who start with a high wage. The fact that the effects of w_{it} persist for four or five years is indicative that a pure search explanation is unlikely to be the only model. Unless offers are received very infrequently, one would expect the search mechanism to phase out the effects of the initial allocation reasonably quickly.

An indicator of whether state dependence is solely caused by search considerations is available when one looks at people who have been involuntarily separated from their previous job. In a search framework, the quality of job obtained influences future job quality because there is an asset value to the current match. If a worker is involuntarily terminated from his current match, this asset value becomes zero. This reasoning suggests the following test for whether the search story alone can explain state dependence. Take only workers who start a new job after an involuntary separation from their old job. Estimate how the wage on their new job depends on the wage on their old job. If there is state dependence in the wages of these workers, it implies that search considerations alone do not explain all state dependence in job quality.[16]

I take all cases where a job starts and pick out cases where the worker reports it ends involuntarily. For these cases, I regress the first wage on the subsequent job on the last wage on the job started at t. As before, I use the starting unemployment rate as the instrument for the wage on the job started at t. For the 986 such cases, the estimated coefficient on the wage is 0.637 (0.209). This suggests that the wage on jobs started subsequent to an involunary termination from the previous job are strongly dependent on the last wage on the previous job. This implies that the type of job a worker holds may change the ability of the worker or the market's perception of the ability of the worker.[17] In the next sections, I examine the evidence for these explanations.

Testing For Stigma

In a stigma model, the type of job a searching worker obtains is influenced both by noisy signals of ability and by exogenous shocks. Since the market does not know the value of realizations of past values of the shock, workers who receive poor draws, and so get bad jobs, will appear to have lower ability than they really have. Therefore, workers who exogenously attain a poor job are stigmatized if potential employers use the wage as a signal of ability. In appendix 2, I formally develop a stigma model. The implication of the model that I use to test for stigma is that over time in a match, the employer learns more about the worker's productivity and so places less weight on the last wage on the previous job as a signal of the worker's ability. My approach is analogous to that of Farber and Gibbons (1996) and Altonji and Pierret (2001) who examine whether employers use education as a signal of ability.

As shown in proposition 1 in appendix 2, the stigma model has the following implication: As tenure with an employer increases, the interactions of tenure with hard to observe variables that are correlated with productivity have positive coefficients in a wage equation and, when these interactions are included, the interactions of tenure with productivity-enhancing variables that are observed by employers when the match formed should become negative. I assume that

[16] In my sample, 63 percent of the involuntary separations are layoffs. It is possible that some of these involuntary separations are temporary layoffs where the worker expected recall but received a better offer while waiting. In this case the asset value of the match remains for the worker. Other research indicates that about 75 percent of laid off workers expect to be recalled, and 72 percent of workers who expect recall are subsequently recalled (Katz and Meyer, 1990). These numbers would imply that only 13 percent of workers who start jobs with new employers after an involuntary separation expected to be recalled by their previous employer.

[17] It could also be the case that the wage on the previous job affects the reservation wage for some irrational reason. For example, the worker may be unwilling to accept a lower wage than the one in the previous job because of pride.

the last wage on the previous job is observed by the employer when the worker is hired. I assume that father's education, mother's education, and the final wage on the job previous to the previous job are not fully observed by the employer when the worker is hired. However, these variables are correlated with productivity on the job and, hence, the signals the employer sees as the worker acquires tenure. The stigma model implies that when interactions of these variables with tenure are included in a wage equation, the interaction of the previous wage with tenure should become negative.[18]

The empirical specification is that the current wage is a function of the last wage on the previous job, fathers and mother's education, tenure, and interactions of tenure with previous wage and with parent's education. I also include controls for tenure squared, a cubic in experience, indicators for married, high school graduate, college graduate, white, a quadratic time trend, and the state indicator variables. Because tenure is correlated with both worker quality and match quality, the OLS coefficients may suffer from heterogeneity bias. Therefore, I also estimate an equation that includes match fixed effects.[19] The identification in this specification is solely from changes over time within employer-employee matches.

The results are in tables 5 and 6. I concentrate on the match fixed effects results because they are robust to unobserved individual and match heterogeneity. In column 4 of table 5, the only tenure interaction is the interaction of tenure with the previous wage. The coefficient on the interaction is -0.0014 (0.0017). The stigma model implies that, with no other interactions, the coefficient on the interaction of tenure with the last wage on the previous job should be close to zero. The reasoning is that if employers use the information in the previous wage efficiently in setting the first wage on the job, subsequent wage changes should not be correlated with the last wage on the previous job.

In column 2 of table 5, I report the coefficients when the only tenure interactions are with parents' education. These interactions have positive coefficients in keeping with the notion that they are correlated with information that is revealed as the match progresses. When, in column 6, all three interactions are included, the coefficient on the previous wage interaction becomes negative and statistically significant. These results are consistent with the predictions of the

---

[18] The econometric implications of statistical discrimination models like the stigma model have been explored by Altonji and Pierret (1995). Altonji and Pierret show that if there is a productivity-enhancing variable, s, that is observed by employers and a variable, z, that is unobserved by employers but correlated with productivity, the coefficient on the interaction of s with time is likely to be negative and the coefficient on the interaction of z with time positive. Their paper explores the effects of a public information model and their empirical specifications interact s and z with experience in the labor market.

[19] There is a sizeable empirical literature that stresses the importance of taking account of individual and match heterogeneity when estimating the return to tenure. See, for example, Topel (1991) and Altonji and Shakotko (1987).

stigma model – the previous wage becomes less important in determining wages as more information about the worker is acquired.

The results in table 5 are also consistent with a human capital explanation whereby ability proxied by parent's education is more closely related to human capital accumulation than the last wage on the previous job. Thus, over time on the job, parents' education becomes more important in determining wages.[20] For this reason, I now carry out the analysis using the last wage on the job previous to the previous job as the partially observed variable. It is unlikely that this wage is more correlated with human capital accumulation than the last wage on the previous job. The results are presented in table 6.[21] Once again, I focus on the results from the match fixed effects specification. As predicted by the stigma model, the coefficient on the last wage on the job previous to the previous job interacted with tenure is positive and when this interaction is added to the wage equation, the interaction of the last wage on the previous job with tenure becomes significantly negative. The coefficients here are of some economic significance. The coefficients of -0.008 and 0.008 in column 6 of table 8 imply that if a worker accumulates 5 years of tenure, the weight on the old information falls by 0.04 and the weight on the new information rises by 0.04. This indicates a substantial reweighting from old to new information about worker ability over the course of the match. The results suggest that employers do use the last wage on the previous job as a signal of ability and that this signal becomes less important as new information is acquired about the employee.

Testing the Learning by Doing Explanation

If a worker exogenously gets a job that is better than his ability would suggest, he may increase his ability through learning by doing or other human capital accumulation mechanisms. Even if workers must pay for learning opportunities, it is possible for a worker to obtain a job that involves both a higher wage and greater learning opportunities. For example, if a worker finds a good match in which he is very productive, he may receive a higher wage than if he had found a poor match even after paying for the greater learning by doing opportunities available. [22]

[20] This possibility is supported by the fact that when I add the interaction of education and tenure to the wage equation, the previous wage interaction becomes more negative, and the coefficients on the parent's education interactions become smaller. The coefficient of the education interaction itself is strongly positive, and it is larger in size than the coefficients on the parents' education interactions.

[21] There are fewer observations used in this table because an individual must have started at least two jobs during the sample period in order to have information on the last wage in the job previous to the previous job.

[22] There is a substantial literature that studies the incidence on training. The findings indicate that the incidence and duration of training is positively related to firm size and to starting wages. (See Altonji and Spletzer ,1991, Bronars and Famulari, 1994).

In 1976, 1978, and 1985, respondents were asked the following question:

"On a job like yours, how long would it take the average new person to become fully trained and qualified?"[23]

I assume that jobs in which the average new person takes a lot of time to become fully trained and qualified are jobs that offer a lot of opportunities for learning by doing. This variable has been widely used as a measure of learning by doing and as a proxy for training (Duncan and Hoffman, 1979; Brown, 1989; Mincer, 1988). Similar questions in other data sets have been used as a training measure (Barron, Black, and Loewenstein, 1993). It is generally found that this variable and other training variables are positively related to wages. However, high-paying jobs will tend to be held by high ability workers so it is not clear that if a worker gets a job that pays better than his usual job that there are more opportunities for learning by doing. This is the hypothesis I test in this section.

Because the training question is asked only in 1976, 1978, and 1985, there are no observations with this information for jobs that start after 1985. If a job is observed in one of these three years, I attach the value reported in that year to all years of the job. For example, if a person starts a job in 1982 and is still in the same job in 1985, I attach the answer to the question in 1985 to the observation in 1982 and all other years of the job. If I have multiple answers to this question for a job because it is observed in at least two of the three years, I take the mean of the answers and use this value in each period of the job. Thus, I have information on learning by doing opportunities for many jobs that do not start in any of the three periods.[24]

I test whether conditional on ability, workers get more learning by doing opportunities if they get a high-paying job. Thus, for the sample of job starters for whom the learning by doing information is available, I regress the log of months of required training (LBD) on the starting wage rate and on a set of worker characteristics:

$$LBD_{it} = \beta_1 w_{it} + \beta_2' x_{it} + f_i + v_{it} \qquad (2)$$

I have experimented with different functional forms for LBD such as using months rather than log(months) and the qualitative result is robust to functional form. Once again, the main problem in estimating this equation is unobserved heterogeneity that is positively correlated with

[23] For other work based on this question, see Brown (1989), and Mincer (1988).

the wage. Unlike in the state dependence estimation, fixed effects is consistent here because there is no lagged dependent variable. Therefore, this is the estimation method I use to estimate equation (2). I have also tried using the state unemployment rate as an instrument for this estimation but because the data is only available for a limited number of years, there is no first stage.

The results of the learning by doing data analysis are in table 7. The sample is composed of new hires only. In table 10, I test how the amount of required training on the job depends on the starting wage. In column 1, I present OLS results for the regression of months of required training on the first wage on the job and worker characteristics. The results indicate that jobs that pay well also provide greater opportunities for learning. The significant coefficients on some of the worker characteristics indicate that these characteristics have large effects on the amount of required training. Thus, it is likely that the coefficient on the starting wage is biased upwards because of unobserved worker heterogeneity. Therefore, in column 2, I present fixed effects estimates of the equation. The coefficient on the starting wage falls somewhat from column 1 but is still large and statistically significant. It does appear that if a worker attains a job with a high starting wage, he also has greater opportunities to learn new skills.[25]

One remaining issue is whether the skills acquired are specific to the job or are more general in nature. I have verified that taking a job with a lot of required training increases wages on subsequent jobs (results available on request). Thus, when workers get a good job with good learning by doing opportunities, at least some of the skills acquired are useful on other jobs.[26] The evidence indicates that learning by doing explanations may be relevant to explaining state dependence in wages.

6. CONCLUSIONS

In this paper, I have used various instrumental variables estimators to determine the extent of state dependence in wages and job quality. My conclusion is that there is a substantial amount of state dependence in wages. If two identical individuals start jobs that pay a different wage, about 60% of the wage differential is still present four years later. Also, approximately

---

[24] Using information on only jobs that start in one of these three years gives similar but less precise results.
[25] It is possible that wages are high on jobs that offer greater learning opportunities because the training increases productivity and wages. By using only the first reported wage on the job in the analysis, I ensure that it is unlikely that this effect is driving the result. In fact, human capital theory would suggest that wages should be temporarily low during training, especially if the training is general.
[26] Brown (1983), Duncan and Hoffman (1979), and Mincer (1988) also find, using the same variable, that learning by doing opportunities have a positive effect on wage growth.

50% of the wage differential is still present even for men who change employers during the four years. The results indicate that the labor market that matches workers to jobs is far removed from the frictionless perfectly competitive model. In practice, the job a searching worker attains is influenced by factors other than skills and talents. Most importantly, this randomness influences the job holdings and wages of individuals to a significant extent for a long period of time.

I have also examined the roles played by different models of state dependence. The first conclusion is that state dependence in wages does not purely arise because jobs persist. The second conclusion is that state dependence goes beyond what would be predicted by a simple search model in which one's wage affects one's reservation wage but not the wage offer distribution faced. Therefore, I have considered evidence for stigma and learning for doing explanations. The evidence suggests that both of these explanations have some relevance. Workers who attain better jobs can increase their skills and hence their future job offers. This implies that human capital models should place more emphasis on the differential opportunities to acquire human capital that exist in different jobs. It also indicates the importance of retraining for workers who have been displaced to less-skilled jobs. The evidence also suggests that taking a bad job is a negative signal to the market. Thus, exogenously attaining a job that is poor relative to ability stigmatizes workers and makes it more difficult for them to get good jobs and high wages in future. As such, the results add to the growing empirical literature that indicates that information models are relevant to understanding labor market dynamics.

REFERENCES

Abowd, J.M., and Card D. (1989), "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, 57, 411-445.

Altonji J. and C. Pierret. (2001) "Employer learning and statistical discrimination", *Quarterly Journal of Economics*, February, 1-37.

Altonji J. and R. Shakotko. (1987), "Do Wages rise with Job Seniority?", *Review of Economic Studies*, 54, 437-459.

Altonji J. and Spletzer, J. (1991), "Worker Characteristics, Job Characteristics, and the Receipt on On-the-Job Training", *Industrial and Labor Relations Review*, 45, 58-79.

Anderson, T.W., and Hsiao, C. (1982), "Formulation and Estimation of Dynamic Models using Panel Data", *Journal of Econometrics*, 18, 47-82.

Arellano, M. and Bover, O. (1995), "Another Look at the Instrumental Variables Estimation of Error Components Models", *Journal of Econometrics*, 29-51.

Baker, M. (1997), "Growth-rate heterogeneity and the covariance structure of life-cycle earnings", Journal of Labor Economics, 15, 338-75.

Barron, J.M., Black D.A., and Loewenstein, M.A. (1993), "Gender Differences in Training, Capital, and Wages", *Journal of Human Resources*, 28(2), 343-364.

Beaudry P., and DiNardo, J. (1991), "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data", *Journal of Political Economy*, 99, 665-688.

Bound, John, Charles Brown, Greg Duncan, and Willard Rodgers (1994), "Evidence on the validity of cross-sectional and longitudinal labor market data", *Journal of Labor Economics* 12: 345-368.

Bronars S.G., and Famulari, M. (1994), "Employer-Provided Training, Wages, and Capital Investment", NBER Conference on Research in Income and Wealth, Labor Statistics Measurement Issues, December 15 and 16.

Brown J. (1989), "Why Do Wages Increase with Tenure?" *American Economic Review*, 79, 971-91.

Brown J. and Light A. (1992), "Interpreting Panel Data on Job Tenure", *Journal of Labor Economics*, 10, 219-257.

Duncan, G.J., and Hoffman, S. (1979), "On-the-Job Training and Earnings Differences by Race and Sex", *Review of Economics and Statistics*, 61(4), 593-603.

Farber, H.S., and Gibbons R. (1996), "Learning and Wage Dynamics", *Quarterly Journal of Economics*, 1007-1047.

Fitzgerald, J., Gottschalk, P., and Moffitt R. (1998), "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics", NBER Technical Working Paper #220.

Geweke, J. and Keane, M., (1997), “An Empirical Analysis of Income Dynamics Among Men in the PSID: 1968-1989”, University of Wisconsin-Madison Institute for Research on Poverty Discussion Paper # 1127-97.

Hause, J.C. (1980), “The Fine Structure of Earnings and the On-the-Job Training Hypothesis”, *Econometrica*, 48, 1013-1029.

Holtz-Eakin, D., Newey, W., and Rosen H. (1988), “Estimating Vector Autoregressions Using Panel Data”, *Econometrica*, 1371-1395.

Katz, L.F. and Meyer, B.D. (1990), “Unemployment Insurance, Recall Expectations, and Unemployment Outcomes”, *Quarterly Journal of Economics*, CV, 973-1002.

Kearl, J.R. (1988), “The Covariance Structure of Earnings and Income, Compensatory Behaviour, and On-the-Job Investments”, *Review of Economics and Statistics*, 70, 214-223.

Little, R.J.A., and D.B. Rubin (1987), *Statistical Analysis with Missing Data*, New York: Wiley.

Ma, C. and Weiss, A. (1993), “A signaling Theory of Unemployment”, *European Economic Review*, 135-57.

McCormick, B. (1990), “A Theory of Signalling During Job Search, Employment Efficiency, and ‘Stigmatised’ Jobs”, *Review of Economic Studies*, 57, 299-313.

MaCurdy, T. (1982), “The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis”, *Journal of Econometrics*, 18, 83-114.

Mincer, J. (1988), “Job Training, Wage Growth, and Labor Turnover”, NBER Working Paper 2690, August.

Mortensen D.T. (1986), “Job Search and Labor Market Analysis”, *Handbook of Labor Economics*, 849-919.

Nickell, S. (1981), “Biases in Dynamic Models with Fixed Effects”, *Econometrica*, 49, 1417-1426.

Parent, D. (1995), “Matching, Human Capital, and the Covariance Structure of Earnings”, Working Paper #351, Industrial Relations Section, Princeton University.

Solon, Gary, Barsky, Robert, and Jonathan A. Parker, 1994, “Measuring the Cyclicality of Real Wages: How Important is Composition Bias?”, *Quarterly Journal of Economics*, 1-26.

Topel R. (1991) "Specific Capital, Mobility, and Wages: Wages rise with Job Seniority", *Journal of Political Economy*, 99, 145-176.

Topel R., and Ward M. (1992), “Job Mobility and the Careers of Young Men*”, Quarterly Journal of Economics*, 439-79.

Waldman, M. (1984), “Job Assignment, Signaling, and Efficiency”, *Rand Journal of Economics*, 15, 255-67.

Table 1: Means of Selected Variables: PSID 1976-1992 (Standard Deviations in Parenthesis)

	(1)	(2)
Variable	All Employed	New Hires
Years of Education	12.80	12.66
	(2.48)	(2.40)
University Degree	0.26	0.22
	(0.44)	(0.42)
High School Diploma	0.51	0.51
	(0.50)	(0.50)
Mother's Education	11.02	11.05
	(2.86)	(2.77)
White	0.69	0.66
	(0.46)	(0.48)
Father's Education	10.54	10.56
	(3.34)	(3.36)
Years of Experience	13.93	11.69
	(8.15)	(7.59)
Government	0.18	0.12
	(0.38)	(0.33)
Salaried	0.41	0.32
	(0.49)	(0.48)
Months of Tenure with Employer	47.47	5.81
	(56.03)	(3.66)
New Job	0.29	1
	(0.45)	(0)
Log of Average Hourly Earnings	2.29	2.09
	(0.56)	(0.57)
Log wage	2.20	2.04
	(0.50)	(0.51)
Predicted Wage	2.11	2.04
	(0.31)	(0.32)
Months to be fully trained	20.29	17.41
	(24.75)	(23.19)
No of Observations	26871	7675

Table 2: Effect of Wage at t on Future Wages For Those Who Start Jobs at t

$$w_{i\tau} = \beta_1 w_{it} + \beta_2' x_{it} + \beta_3 u_\tau + \beta_4 u_{\tau-1} + f_i + \eta_{i\tau}$$

Instrument	1st Stage Coefficient on Instrument	$w_{i\tau} = w_{it+1}$	$w_{i\tau} = w_{it+2}$	$w_{i\tau} = w_{it+3}$	$w_{i\tau} = w_{it+4}$	$w_{i\tau} = w_{it+5}$
OLS		0.705 (0.014) [5668]	0.658 (0.015) [4953]	0.613 (0.019) [4327]	0.562 (0.021) [3758]	0.575 (0.023) [3255]
$w_{it} - w_{it-1}$	0.480 (0.018)	0.408 (0.036) [4347]	0.314 (0.036) [3797]	0.272 (0.043) [3293]	0.187 (0.045) [2873]	0.184 (0.047) [2461]
$w_{it-1} - w_{it-2}$	0.189 (0.033)	0.890 (0.179) [2747]	0.825 (0.177) [2393]	0.604 (0.198) [2088]	0.616 (0.176) [1823]	0.543 (0.220) [1534]
$e_{it} - e_{it-1}$	0.101 (0.015)	1.576 (0.189) [4554]	0.910 (0.136) [3971]	0.953 (0.151) [3432]	0.646 (0.137) [2987]	0.573 (0.139) [2566]
SU_t	-0.023 (0.004)	0.439 (0.149) [5668]	0.519 (0.121) [4953]	0.705 (0.157) [4327]	0.610 (0.175) [3758]	0.583 (0.224) [3255]

The estimates in the table are the coefficients on w_{it} from each specification.
Also included in each specification are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, and a quadratic time trend.
Robust standard errors in parentheses account for repeated observations on individuals.
The first stage estimates reported are from the t+4 regressions. The first stages in other years are similar.
SU_t is the starting state unemployment rate of the job started at t.
The number of observations used in each specification is in square brackets.

Table 3: Effect of Predicted Wage at t on Future Predicted Wages For Those Who Start Jobs at t

$$pw_{i\tau} = \beta_1 pw_{it} + \beta_2^{/} x_{it} + \beta_3 u_{\tau} + \beta_4 u_{\tau-1} + f_i + \eta_{i\tau}$$

Instrument	1st Stage Coefficient on Instrument	$pw_{i\tau} = pw_{it+1}$	$pw_{i\tau} = pw_{it+2}$	$pw_{i\tau} = pw_{it+3}$	$pw_{i\tau} = pw_{it+4}$	$pw_{i\tau} = pw_{it+5}$
OLS		0.626 (0.013) [5668]	0.516 (0.016) [4953]	0.496 (0.017) [4327]	0.467 (0.190) [3758]	0.435 (0.022) [3255]
$pw_{it} - pw_{it-1}$	0.461 (0.017)	0.350 (0.033) [4347]	0.222 (0.036) [3797]	0.157 (0.039) [3293]	0.176 (0.042) [2873]	0.181 (0.045) [2461]
$pw_{it-1} - pw_{it-2}$	0.070 (0.025)	0.982 (0.285) [2747]	0.847 (0.304) [2393]	0.811 (0.329) [2088]	0.758 (0.384) [1823]	0.486 (0.351) [1534]
$e_{it} - e_{it-1}$	0.038 (0.010)	1.017 (0.252) [4554]	0.263 (0.220) [3971]	0.890 (0.397) [3432]	0.687 (0.264) [2987]	0.737 (0.348) [2566]
SU_t	-0.014 (0.003)	No First Stage	0.673 (0.207) [4953]	0.608 (0.216) [4327]	0.483 (0.195) [3758]	0.019 (0.246) [3255]

The estimates in the table are the coefficients on pw_{it} (the predicted wage) from each specification.
Also included in each specification are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, and a quadratic time trend.
Robust standard errors in parentheses account for repeated observations on individuals.
The first stage estimates reported are from the t+4 regressions. The first stages in other years are similar.
SU_t is the starting unemployment rate of the job started at t.
The number of observations used in each specification is in square brackets.

Table 4: Effect of Wage at t on Future Wages For Persons Who Start Jobs at t (Specification Checks for starting Unemployment Rate Instrument)

Specification Check	w_{it+1}	w_{it+2}	w_{it+3}	w_{it+4}	w_{it+5}
Adding interactions of personal characteristics with the quadratic time trend	0.405 (0.162) [5668]	0.499 (0.125) [4953]	0.690 (0.164) [4327]	0.591 (0.180) [3758]	0.571 (0.231) [3255]
Adding Year Indicators	0.717 (0.203) [5668]	0.634 (0.244) [4953]	0.747 (0.263) [4327]	0.668 (0.304) [3758]	0.862 (0.355) [3255]
Adding controls for whether job started as a result of a voluntary or involuntary termination	0.424 (0.157) [5668]	0.509 (0.135) [4953]	0.737 (0.181) [4327]	0.647 (0.202) [3758]	0.579 (0.273) [3255]
Adding Controls for the mean state unemployment rate while the individual is employed and the mean starting unemployment rate of the individual	0.526 (0.148) [5668]	0.612 (0.116) [4953]	0.714 (0.143) [4327]	0.535 (0.173) [3758]	0.497 (0.210) [3255]
Excluding Poverty Sample	0.653 (0.168) [3169]	0.511 (0.172) [2822]	0.543 (0.245) [2495]	0.744 (0.258) [2187]	0.611 (0.317) [1925]
Maintaining Constant Sample in each year	0.427 (0.277) [2139]	0.411 (0.187) [2139]	0.778 (0.212) [2139]	0.762 (0.230) [2139]	0.656 (0.287) [2139]
Weighting for Attrition	0.380 (0.161) [5668]	0.543 (0.120) [4953]	0.688 (0.159) [4327]	0.589 (0.170) [3758]	0.498 (0.219) [3255]
Weighting for Attrition and Missing Wages	0.398 (0.156) [5668]	0.521 (0.125) [4953]	0.651 (0.164) [4327]	0.533 (0.180) [3758]	0.496 (0.226) [3255]
Replacing missing wages with average hourly earnings	0.386 (0.186) [6646]	0.478 (0.137) [5907]	0.701 (0.173) [5233]	0.689 (0.197) [4611]	0.418 (0.266) [4045]

The estimates in the table are the coefficients on w_{it} (the predicted wage) from each specification.
Also included in each specification are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, and a quadratic time trend.
Robust standard errors in parentheses account for repeated observations on individuals.
All specifications are estimated by 2SLS with the starting unemployment rate used as the instrument.
The number of observations used in each specification is in square brackets.

Table 5: The Effects of Wage on Previous Job and Parents' Schooling on Wages: Ordinary Least Squares and Match Fixed Effects

	Ordinary Least Squares			Match Fixed Effects		
Observations: 11550 (2803 Matches)	(1)	(2)	(3)	(4)	(5)	(6)
Previous Wage	0.551 (0.020)	0.524 (0.021)	0.566 (0.021)			
Mother's Education	0.0036 (0.0028)	-0.0026 (0.0027)	-0.0034 (0.0031)			
Father's Education	0.0069 (0.0024)	0.0028 (0.0027)	0.0022 (0.0027)			
Previous Wage * Tenure(years)	-0.0089 (0.0044)		-0.0134 (0.0044)	-0.0014 (0.0017)		-0.0043 (0.0018)
Mother's Education * Tenure(years)		0.0019 (0.0007)	0.0021 (0.0007)		0.0010 (0.0003)	0.0011 (0.0003)
Father's Education* Tenure(years)		0.0014 (0.0007)	0.0016 (0.0007)		0.0009 (0.0003)	0.0010 (0.0003)

Also included in the equation are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, a quadratic time trend, and a quadratic in tenure.
Robust standard errors in parentheses account for repeated observations on individuals.
Matches that only last for one period are omitted.

Table 6: The Effects of Wage on Previous Job and the Wage on Job Previous to Previous Job on Wages: Ordinary Least Squares and Match Fixed Effects

	Ordinary Least Squares			Match Fixed Effects		
Observations: 7276 1831 Matches	(1)	(2)	(3)	(4)	(5)	(6)
Previous Wage	0.441 (0.028)	0.389 (0.033)	0.443 (0.030)			
Lagged Previous Wage	0.232 (0.029)	0.266 (0.029)	0.229 (0.031)			
Previous Wage * Tenure(years)	-0.017 (0.007)		-0.018 (0.012)	-0.0034 (0.0020)		-0.0076 (0.0026)
Lagged Previous Wage * Tenure(years)		-0.011 (0.007)	0.001 (0.012)		0.0023 (0.0023)	0.0076 (0.0029)

Also included in the equation are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, a quadratic time trend, and a quadratic in tenure.
Robust standard errors in parentheses account for repeated observations on individuals.
Matches that only last for one period are omitted.

Table 7: Determinants of Length of Time Required for Average Worker to Become Qualified, New Hires

Dependent Variable: Log (Time Required)

Variable	OLS	Fixed Effects
Log (Wage)	1.125 (0.083)	0.524 (0.206)
Married	0.170 (0.074)	0.145 (0.205)
Years of Experience	0.021 (0.034)	0.030 (0.257)
Experience Squared/100	-0.008 (0.232)	-0.696 (0.491)
Experience Cubed/10000	-0.071 (0.452)	1.431 (0.917)
College Degree	0.442 (0.097)	
High School Graduate	0.137 (0.082)	
White	0.504 (0.080)	

There are 1929 observations on 1536 people.
Also included in the equation are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, a quadratic time trend, and a quadratic in tenure.
Robust standard errors in parentheses account for repeated observations on individuals.

Appendix 1: Sample Attrition

Because of attrition and non-employment many future wages are unobserved. The issues are similar in all the regressions; for exposition purposes I detail the situation for the wage at t+4. While there are 7675 observations on job starts at time t with valid wages, there are only 3758 observations at t+4 with non-missing wages and non-missing observations on the other variables used in the analysis. There are several reasons for the non-availability of these 3917 cases. First, 1930 cases are missing because the job starts after 1988 and hence we do not follow people long enough to observe the wage at t+4. Second, the attrition of subjects from the PSID survey is responsible for 884 of the missing cases. Third, of those that remain in the survey, 716 are non-employed at the survey date and hence have missing wage data. Fourth, 387 people are employed at the survey period at t+4 but have missing wage data. For most of these people, wage data is missing because they are not paid either hourly or salaried.

Obviously, there are selection issues that arise because the data is non-missing at t+4 in a non-random fashion. It is well known that the probability of attriting from the PSID is related to the observable characteristics of workers and to their wages at t. (for example, see Fitzgerald, Gottschalk, and Moffitt, 1998). Likewise, the probability of not working at t+4 may depend on wages at t. Therefore, this issue requires some attention.

A formal approach to the missing data problem is to model the probability that a worker will have non-missing data at t+4. I allow the probability of attrition to depend on the variables that are always observed, w_{it} and x_{it} in this case, but I do not allow dependence on the variable that is missing for some units, $w_{i,t+4}$. Let D_i be an indicator function that takes the value of 1 if there is full data for an individual at t+4 and zero otherwise.

$$\Pr(D_i = 1 \mid W_{it} = w_{it}, W_{i,t+4} = w_{i,t+4}, X_{it} = x_{it}) = g(w_{it}, x_{it}) \qquad \text{(A1)}$$

with g(.) unknown. The estimation of this model is carried out in two steps. In the first stage, I estimate a probit model that conditions the probability of attrition on w_{it} and x_{it}.[27] The predicted probabilities from the probit model are used to form weights and these weights are used to weight the observations in the 2SLS estimation in the second step. The weights are equal to the inverse of the probability of not attriting between t and t+4.

This is referred to as Missing at Random (Little and Rubin, 1987). The argument is that there is nothing in the data that suggests that units that drop out are systematically different from units who do not drop out once we condition on all observed variables. This model has some intuitive appeal. Consider a unit that drops out between t and t+4 with values of the observed variables equal to $W_{it} = w_{it} \text{ and } X_{it} = x_{it}$. The Missing at Random assumption implies that for our best guess of the value of the missing variable $W_{i,t+4}$, we should look at values of $W_{i,t+4}$ for units with the exact same values of w_{it} and x_{it}.

The results of the estimation are contained in table A1. Column 1 of table A1 contains the derivatives from probit estimation of the probability that a person does not attrit from the PSID between t and t+4. The probability of attriting is lower for people with higher wages at t, persons with college degrees, married people, and white people. In the second column of table A1, I report the derivatives from probit estimation of the probability that all relevant data is present at t+4 given it is present at t. Clearly, persons with missing data at t+4 differ from persons with complete data. Observations present at t+4 have higher wages at t, are more likely to have a college degree at t, are more likely to be married at t, and are more likely to be white.

[27] I allow the latent variable to depend on the explanatory variables in a linear fashion in the probit estimation. However, adding interactions of the wage at t with the x variables and other non-linearities does not change the weighted 2SLS results to any appreciable extent.

Table A1: Probit Estimates of Probability Worker who starts Job at t has valid data at t+4

	(1)	(2)
	Probability still PSID respondent at t+4 (Probit 1)	Probability employed at t+4 with valid observations on wage and other covariates (Probit 2)
Number of Observations:	Derivatives	Derivatives
W_{it}	0.034 (0.011)	0.077 (0.018)
College Degree	0.040 (0.014)	0.126 (0.022)
High School Graduate	-0.042 (0.012)	0.018 (0.019)
Married	0.022 (0.010)	0.031 (0.018)
State Unemployment Rate	0.004 (0.003)	0.010 (0.004)
White	0.059 (0.013)	0.090 (0.020)
Experience at t	0.001 (0.006)	0.004 (0.008)
Experience at t Squared/100	-0.001 (0.038)	-0.028 (0.055)
Experience at t Cubed/10000	0.001 (0.073)	0.048 (0.103)

Also included in the equation are experience, experience squared, experience cubed, indicators for college degree, high school graduate, married, and white, the state unemployment rate at τ, the state unemployment rate at τ-1, state dummies, a quadratic time trend, and a quadratic in tenure.
Robust standard errors in parentheses account for repeated observations on individuals

Appendix 2: A Stigma Model

In this model, the type of job a searching worker obtains is influenced both by noisy signals of ability and by exogenous shocks. Since the market does not know the value of realizations of past values of the shock, workers who receive poor draws and get bad jobs will appear to have lower ability than they really have. Therefore workers who exogenously attain a poor job are stigmatized if employers use job quality as a signal of ability.

Assumptions

(i) There is a group of workers who vary in terms of ability (a) but are observationally equivalent. All these workers look for a job at time one.

(ii) Each period the market sees a noisy signal (θ) of ability. This signal has the following distribution: $\theta \sim N(a, 1/r)$. The signal is iid.

(iii) A searching worker gets a job of quality $w_t = E(a | I_t) + \varepsilon_t$ where I_t is the information set available to the market at t.[28] I define a job as a match with an employer. Except in period 1, when there is no previous wage to be observed, the market's information set at t consists of the wage at t-1 and the signal emitted at t. The market is assumed to use Bayes' law to update beliefs about the workers' abilities. The noise in the wage allocation process reflects luck that causes workers to get jobs that over or understate their ability. I assume that $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

(iv) The current employer sees a normally distributed signal, η, of the worker's ability each period. I assume that η is more precise than θ because the employer has an informational advantage over the market. The precision of η is k. The employer's information set at t is Ω_t. This information set contains the last wage on the worker's previous job, the value of θ that the market observed when the worker was hired, and all the realizations of η since the worker started the job.

(v) Workers can quit to another job if they get a better offer. Also, matches end exogenously with a certain probability.

[28] For simplicity, I assume that workers accept this offer. It would be optimal for workers to accept an offer if the present value of doing so exceeds the present value of waiting for a better offer. Since poor jobs are stigmatizing, sometimes it will be optimal to wait.

(vi) The offer to a worker from his current employer is $w_t^0 = \alpha w_{t-1} + (1-\alpha)E(a|\Omega_t) + v_t$. The value of α reflects the extent to which the employer's wage offer adjusts to his new information about worker ability.[29] I assume that α is an exogenous constant.

Proposition I

(a) There is state dependence in job quality. In other words, a worker's wage at t depends on his wage at t-1. A poor draw of ε_1 reduces w_1 and thus reduces w_2 because $w_2 = \Pi_1.w_1 + \Pi_2.\theta_2 + \varepsilon_2$; where $\Pi_1, \Pi_2 > 0$.

(b) As tenure on any job increases, the employer places less weight on the last wage on the previous job. Also less weight is placed on the realization of θ at the start of the job. Increased weight is placed on realizations of η as these are observed over time.

The proof of (a) is as follows. At time one, the market determines the expected value of the workers' ability and sets w_1. At time two, the market sees w_1 and θ_2. The market updates its expectation of each worker's ability in a bayesian fashion. Thus the posterior expectation is

$$E(a|w_1,\theta_2) = \left[\frac{1}{2+r\sigma_\varepsilon^2}\right].w_1 + \left[\frac{1+r\sigma_\varepsilon^2}{2+r\sigma_\varepsilon^2}\right].\theta_2 = \Pi_1.w_1 + \Pi_2.\theta_2$$

Competition implies that if the worker changes employer, $w_2 = \Pi_1.w_1 + \Pi_2.\theta_2 + \varepsilon_2$

The current employer sees a normally distributed signal, η_2, of the workers ability. The employer's information set is $\{\theta_1, \eta_2\}$. Then, the employer's estimate of the ability of the worker is

$$E(a|\eta_2,\theta_1) = \left[\frac{r\theta_1 + k\eta_2}{r+k}\right].$$

The employer's offer is $w_2^0 = \alpha w_1 + (1-\alpha)E(a|\eta_2,\theta_1) + v_2$.

Therefore, whether the worker separates from the job or not, the wage at time two depends on the wage at time one. One can see that a poor draw of ε_1 stigmatizes the worker in the sense that the market's expectation of his ability is less than his true ability and less than his employer's estimate of his ability.

---

[29] Since η is not observed by the market, employers may not have to raise wages for workers with high realizations of η. However, employers will want to place more capable employees into positions of responsibility (Waldman, 1984). If the market infers ability from the position of the worker, the employer will be forced to raise wages for workers who have high values of η.

To prove part (b), note that in period three the market receives the signal θ_3, but the employer instead sees the signal η_3. Assume the worker started his current job at time 2. The employer's information set is now $\{w_1, \theta_2, \eta_3\}$. The information set of the outside market is $\{w_2, \theta_3\}$. The employer and the market both update their beliefs about the worker's ability. The precision of the posterior distribution of a when the worker begins a new job is just the sum of the precision of the prior and the precision of the signal. Thus,

$$\Pr ecision(E(a|\, w_1, \theta_2)) = \left[\frac{r}{1+\sigma_\varepsilon^2 r}\right] + r = \left[\frac{2r+\sigma_\varepsilon^2 r^2}{1+\sigma_\varepsilon^2 r}\right] = l\,.$$

The bayesian update of the employer's belief is

$$E(a|\, w_1, \theta_2, \eta_3) = \left[\frac{l}{k+l}\right]\Pi_1 w_1 + \left[\frac{l}{k+l}\right]\Pi_2 \theta_2 + \left[\frac{k}{k+l}\right]\eta_3\,.$$

The employer's offer is $w_3^0 = \alpha w_2 + (1-\alpha)E(a|\,\eta_3, \theta_2, w_1) + v_3$. One can easily verify that compared to $w_2 = \Pi_1 . w_1 + \Pi_2 . \theta_2 + \varepsilon_2$, the weights on w_1 and θ_2 have fallen, and the weight on η_3 is now positive. Thus, the weights on the signals observed at the start of the job fall with tenure on the job and, as tenure increases, greater weight is placed on variables that are correlated with productivity but not fully reflected in the initial signal and in the previous wage.