

BLS WORKING PAPERS



U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Employment and Unemployment Statistics

Wage Compression and the Division of Returns to Productivity Growth: Evidence from EOPP

Harley Frazis, U.S. Bureau of Labor Statistics
Mark A. Loewenstein, U.S. Bureau of Labor Statistics

Working Paper 398
August 2006

All views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics.

Wage Compression and the Division of Returns to Productivity Growth: Evidence from EOPP

Harley Frazis
Bureau of Labor Statistics
2 Massachusetts Ave. NE
Suite 4945
Washington D.C. 20212

Frazis_H@BLS.GOV

Mark A. Loewenstein
Bureau of Labor Statistics
2 Massachusetts Ave. NE
Suite 4130
Washington D.C. 20212

Loewenstein_M@BLS.GOV

August 2006

Any views expressed in this paper are the authors' and do not necessarily reflect those of the Bureau of Labor Statistics.

I. Introduction

The relationship between wages and productivity is fundamental to labor economics. In this paper, we examine this relationship using subjective measures of worker productivity from the Employment Opportunity Pilot Project (EOPP) data. We seek to answer the following questions. To what extent are differences in productivity between workers on the same job reflected in their relative wages? To what extent is productivity growth reflected in wage growth? To what extent can data on wage and productivity growth, and the relative wages and productivity of workers on the same job, identify the division of returns to human capital and the extent to which wages are compressed relative to productivity?

We follow Bishop's (1987) paper in examining the relationship between wages and productivity in the EOPP data. Bishop (1987) finds that direct measures of productivity reported by the employer have quite weak effects on wages and notes that indirect indicators of productivity such as relevant experience have stronger effects. We extend Bishop's research by using these indicators as instruments for the noisy measures of productivity.

In the simplest static model of a competitive labor market, the relationship between productivity and wages is straightforward: wages equal marginal product. In dynamic models, such as Becker (1962), this equality need not always hold. In Becker's model, workers bear the full cost and return to acquiring general human capital, so general human capital acquisition does not affect the equality of wages and productivity. In contrast, the costs and returns to acquiring skills specific to the employer are shared between the employer and the worker, so that wages are above marginal productivity before training and below productivity after training. More recent models and empirical work¹ suggest that employers share the costs and returns to general

¹For examples, see Acemoglu and Pischke (1998,1999), Loewenstein and Spletzer (1998), Chang and Wang (1996), Autor (2001), Balmaceda (2005), Bishop (1991), Harhoff and Kane (1997), and Booth and Bryan (2002).

as well as specific human capital. Even if workers realize the full return and bear the full cost of training, market frictions such as search costs (Acemoglu and Pischke (1999) and Zoega and Booth (2005)) may cause wages to diverge from productivity. In addition, as suggested by Akerlof and Yellen (1990) and others, social norms may constrain wages from fully reflecting productivity differentials. We refer to the phenomenon of differences in wages less than fully reflecting differences in productivity, for reasons other than the sharing of the costs of human capital acquisition, as "wage compression".

As detailed below, the EOPP data contain measures of a recently hired worker's productivity at various points in time. Labor economists typically abstract away from the fact that wages are not adjusted continuously. However, our data show that productivity typically grows quite rapidly at the start of a job. Since wages are not adjusted continuously, starting wages very likely reflect not just the low productivity at the very start of the job but also the much greater productivity typically attained within a few weeks. We develop a method that allows us to take this into account in estimating the relationship between starting wages and productivity.

Our data allow us to estimate how productivity growth affects wage growth in the first two years in a job, and how differences in starting productivity between workers affect starting wages. We show that under reasonable assumptions, if one regresses differences in starting wages for workers in the same job on differences in starting productivity and differences in productivity growth, the coefficient on starting productivity identifies the extent to which starting wages are compressed. The coefficient on differences in productivity growth in the same regression identifies the extent to which employers share the cost of general human capital acquisition.

Our results show a dramatic amount of wage compression. Our point estimate from our benchmark specification is that only 32 percent of differences in starting productivity are reflected in differences in starting wages, even after accounting for differences in productivity growth. The hypothesis that starting wages equal starting productivity after taking into account human capital acquisition is decisively rejected. Our results also indicate that productivity growth translates into substantially lower wage growth: productivity growth of 10 percent results in wage growth of only 2.6 percent. This reflects employer sharing of the costs and returns to human capital acquisitions and/or increased wage compression over time. Unfortunately, we do not obtain a precise estimate of the amount of employer sharing of human capital investment per se.

The great extent of wage compression poses a challenge to theories of wage determination. Existing theories of wage compression seem unlikely to explain the magnitude of compression that we find. This is especially true in view of the fact that our wage compression coefficient is to a great extent identified by differences in relevant experience. That employers would pay experienced workers their marginal products seems intuitive and is the basis of Becker's (1962) prediction that employers do not share the costs to general human capital. In our final section we briefly develop a model that generates wage compression via equity norms that prevent employers from paying experienced workers hired from the outside more than workers with the same skills trained at the firm. The “equity norm” not only leads to wage compression, but also amplifies any initial tendency toward employer sharing of the return to general human capital acquisition, as the wage compression and sharing effects reinforce each other.

II. Identifying the Relationship Between Wages and Productivity Using the EOPP Data

EOPP is one of the few surveys that we are aware of that has information on the productivity of individual workers.² The EOPP items on the productivity of the last worker hired and the typical worker at various points in time and the availability of instruments such as a worker's relevant experience elsewhere and job complexity make the data an ideal source for analyzing the relationship between productivity and wages. The EOPP data have been analyzed fairly extensively, but the relationship between wages and productivity has been under-explored. We use the data to estimate the rate at which productivity growth during the first two years of employment translates into wage growth and the extent to which productivity differences among workers who are hired for the same position are reflected in differences in starting wages.

Description of the EOPP Data

The EOPP data come from a 1982 survey sponsored by the National Institute of Education and the National Center for Research in Vocational Education. Employers are asked a series of questions about the wages, experience, productivity, and training activities of the last worker hired and the typical worker in the same position.³ Descriptive statistics can be found in Table 1.

For our purposes, the key EOPP data are the information on wages and productivity. Employers in the survey are asked about the starting wage paid to the last worker hired and the average hourly wage paid to workers who have had two years' experience in the position. In addition, employers in the EOPP survey are asked about the productivity of the last worker hired

² The Small Business Administration survey, which is modeled after EOPP, also has information on worker productivity; see Barron, Black, and Berger (1999). There are a few company data sets with information on productivity. Using performance ratings at a major corporation, Medoff and Abraham (1981) find that that within-grade-level seniority – earnings differentials cannot be explained by performance differentials. Analyzing the personnel records of a large company, Bartel (1995) finds that training is positively related to both wage growth and performance.

³ For more information about the survey and the training questions, see Barron, Black, and Loewenstein (1989).

and the typical worker in the position during the first two weeks of employment and from the “third week to the twelfth week at work.” Employers in the survey also provide an estimate of the productivity of the “typical worker who has been in this job for two years.” In forming their productivity estimates, employers are told to rate workers on a “productivity scale of zero to one hundred, where one hundred equals the maximum productivity any of your employees can attain and zero is absolutely no productivity.”⁴

As mentioned in the introduction, our key extension of Bishop's (1987) research on wages and productivity using these data is to correct for measurement error in productivity by using IV. For IV to produce consistent estimates, measurement error in log productivity must be uncorrelated with true log productivity. One possible objection to our procedure is that the employers answering the EOPP productivity question may tend to exaggerate productivity differences between workers (and also exaggerate productivity growth). Suppose that an employer's productivity estimate is given by

$$\ln p = \alpha_0 + \alpha_1 \ln p^* + e,$$

where p^* is true productivity and e is a mean-zero error uncorrelated with $\ln p^*$. If $\alpha_1 > 1$ (that is, if reported productivity differences are greater in magnitude than actual productivity differences), then measurement error will be positively correlated with $\ln p^*$ and IV estimation of the effect of log productivity on wages will underestimate the coefficient on log productivity.⁵

Examining the data, we find that the average coefficient of variation for productivity for jobs where we have observations for two workers, using observations where the worker has been on the job for more than 0.8 years, is 0.20 (this is the case for both productivity during the first

⁴ Employers also provide information on the productivity and starting wage of the previously most recently hired worker in the same position as the last worker hired, but this information is unfortunately missing in the majority of observations.

⁵ Bollinger and Chandra (2005), writing about general models of measurement error, note that “[t]o our knowledge, no study has ever found $[\alpha_i] > 1$.”

two weeks and productivity from the third to the twelfth week; other tenure cutoffs give similar results). This is almost identical to the average found in a review of studies examining variation in output for which physical measures existed (Schmidt and Hunter 1983), so there is no evidence that our productivity measure exaggerates differences between workers.

We estimate log productivity growth over the first two years as the difference between the typical worker's log productivity after two years and the most recent hire's log productivity after two weeks. As can be seen in Table 1, average productivity growth over the first two years is very high: productivity after two years is on average eighty percent higher than productivity at the start of the job.

We can decompose productivity growth over the two year period into productivity growth occurring at the very start of the employment relationship and productivity growth occurring after the twelfth week. Specifically, we estimate log productivity growth over the first three months as the difference between log productivity from the third to the twelfth week and log productivity during the first two weeks (given the way the question is worded, this may be an underestimate); productivity growth from the twelfth week on is estimated as the difference between log productivity after two years and log productivity from the third to twelfth week. As reported in Table 1, productivity during the third to twelfth week is on average 51 percent higher than productivity during the first two weeks while productivity after two years is 29 percent higher than during the third to twelfth week. Apparently then, productivity does not grow evenly over the first two years, but is heavily concentrated toward the start of the job: at least two-thirds of the growth in productivity over the first two years occurs during the first three months.

Besides the information on productivity and wages, EOPP also contains many additional variables that are useful for our analysis. As alluded to above, EOPP provides fairly detailed

information on both formal and informal training. Employers indicate the number of hours during the first three months of employment that were spent providing formal training through “self-paced learning programs or ... by specially trained personnel” to not only the last worker hired, but also to the typical worker in the same position. With respect to informal training, employers provide information on the number of hours during the first three months of employment that management and line supervisors spent giving individualized training to the last worker hired and the typical worker in the same position. Employers also indicate the number of hours that co-workers who are not supervisors spent away from their normal work giving informal individualized training to the last worker hired and the typical worker. Finally, employers indicate the total number of hours during the first three months of employment that the average new employee in the position spends in training activities watching other people do the job rather than doing it himself.

Two other items in the data are also relevant for our analysis. There is information on the most recently hired worker’s relevant employment experience in jobs “that had some application” to the position for which he was hired. In addition, employers estimate the number of weeks it takes a new employee in the most recently filled position to become fully trained and qualified if he or she has the necessary school provided training but no experience in the job. As discussed by Barron, Black, and Berger (1999) and Frazis and Loewenstein (2005), the latter variable can be thought of as capturing job complexity.

Framework for Empirical Analysis

The EOPP data enable us to estimate two equations: a regression of differences in starting wages on differences in starting productivity for workers in the same job and a regression of wage growth on worker productivity growth. As described above, EOPP has information on

both formal and informal training. As reported in Loewenstein and Spletzer (1999), the employers in EOPP report that most of the skills the new employee learns on the job are useful elsewhere.⁶ Still, we allow for the possibility that some human capital is employer-specific. We assume that human capital acquisition has the multiplicative structure $\ln H = \ln G + \ln S$, where G and S are the worker's stocks of general and specific human capital, respectively, and H is equal to productivity.

Workers will typically have different amounts of general human capital when they start a job, while by definition all workers will have zero firm-specific capital when they are hired. One would therefore expect that *within a job* differences between workers in human capital acquisition will be dominated by differences in general training, even for jobs where there are large amounts of specific training. Indeed, in the case where all workers in a job are brought to the same level of general and specific human capital by the end of a training period, there will be no differences in specific human capital acquisition between workers in the same job. To simplify the discussion below, we assume that specific human capital acquisition does not vary among workers in the same job. We view this assumption as quite reasonable, particularly in light of the fact that there is some question whether much human capital acquisition is specific to start with.

Consider a two period model. Workers' starting wages depend upon their starting productivity and upon their human capital acquisition:

$$(1) \quad \ln w_1 = \alpha_1 + \beta_1 \ln H_1 + \gamma_{1g} g + \gamma_{1s} s + u_1,$$

where H_1 is the worker's stock of general human capital at the start of the match, $g \equiv \ln G_2 - \ln G_1$ and $s \equiv \ln S_2$ are general and specific human capital acquisition during the first period, and u_1

⁶ Seventy-two percent of employers indicate that nearly all or most of the skills learned are general; eight percent indicate that none of the skills are general. Similarly, workers in the NLSY indicate that most of the skills acquired in training are useful at other employers.

is a mean zero residual. In interpreting (1), note that the coefficient β_l denotes the extent of wage compression. If $\beta_l = 1$, differences in productivity are completely reflected in differences in wages once human capital cost sharing is accounted for. If $\beta_l < 1$, differences in wages are compressed relative to differences in productivity. The coefficients γ_g and γ_s are nonnegative and reflect the degree to which the employer shares the cost of human capital acquisition. Cost sharing by the employer means that worker is paid a starting wage higher than he otherwise would receive.⁷ If the worker bears the full cost to general human capital acquisition, then $\gamma_{1g} = 0$.

Similarly, wages in period 2 are given by

$$(2) \quad \ln w_2 = \alpha_2 + \beta_2 \ln H_2 + \gamma_{2g} g + \gamma_{2s} s + u_2.$$

As with equation (1), the coefficient β_2 reflects wage compression. The coefficients γ_{2g} and γ_{2s} are nonpositive and reflect the employer sharing the return to general and specific human capital accumulation, respectively: if the employer shares the return to human capital accumulation, then the worker is paid a wage below what he would otherwise receive.⁸

Estimating (1) and (2) would require measures of productivity in dollar terms and decompositions of measured productivity growth into specific and general components, neither of which is present in the EOPP data. Measurement error is also a concern. In light of these considerations, we use a set of instruments to estimate two differenced equations: a wage growth

⁷ If workers become more productive simply because of learning by doing, then human capital acquisition does not entail an explicit cost. However, as made clear in the model laid out in Section IV, if employers realize part of the return to the higher productivity in period 2, a zero profit equilibrium condition implies that workers will be paid a first period wage that exceeds their starting marginal product, so that γ_{1g} and γ_{1s} are positive.

⁸ Note that the specification in (2) is no less general than the specification

$\ln w_2 = \alpha_2 + \beta_2 \ln G_2 + \beta'_2 \ln S_2 + \gamma_{2g} g + \gamma'_{2s} s + u_2$: the two specifications are equivalent if $\beta_2 + \gamma_{2s} = \beta'_2 + \gamma'_{2s}$. Since $s = \ln S_2$, the sum $\beta'_2 + \gamma'_{2s}$ is identified, but not the parameters β'_2 and γ'_{2s} .

equation that is obtained by subtracting (1) from (2) and a first differenced starting wage equation obtained by differencing (1) between two different workers in the same job.⁹

Differencing the starting wage equation (1) for two workers, A and B, in the same job:

$$(3) \quad \ln w_1^A - \ln w_1^B = \beta_1(\ln H_1^A - \ln H_1^B) + \gamma_g(g^A - g^B) \\ + \gamma_s(s^A - s^B) + (u_1^A - u_1^B) .$$

Consider a regression of $\ln w_1^A - \ln w_1^B$ on predicted differences in the two workers' human capital:

$$(4) \quad \ln w_1^A - \ln w_1^B = b_1 \hat{H}^{A-B} + c_1 \hat{h}^{A-B} + u$$

where $\hat{H}^{A-B} \equiv E((\ln H_1^A - \ln H_1^B) | Z)$,

$$\hat{h}^{A-B} \equiv E((\ln H_2^A - \ln H_1^A) - (\ln H_2^B - \ln H_1^B) | Z) \\ = E((g^A - g^B) + (s^A - s^B) | Z),$$

u is a residual, and Z is a set of instruments. Under our assumption that $(s^A - s^B) = 0$ for all

jobs, $\hat{h}^{A-B} = E((g^A - g^B) | Z)$ and the consistency of 2SLS implies that $p \lim b_1 = \beta_1$ and

$p \lim c_1 = \gamma_g$.¹⁰ In Becker's (1962) classic model of human capital acquisition, there is no wage

⁹ Due to idiosyncrasies of the data which we discuss below, it is not possible to difference second period wages between two workers in the same job.

¹⁰ These results depend on our assumption that workers in the same job acquire the same amount of specific human capital. More generally, project \hat{H}^{A-B} on $\hat{g}^{A-B} \equiv E(g^A - g^B | Z)$ and $\hat{s}^{A-B} \equiv E(s^A - s^B | Z)$:

$\hat{H}^{A-B} = d_0 + d_1 \hat{g}^{A-B} + d_2 \hat{s}^{A-B} + v$. Then one can show that

$$p \lim b_1 = \beta + (\gamma_g - \gamma_s) \frac{(d_1 - d_2)(\text{Var}(\hat{g}^{A-B})\text{Var}(\hat{s}^{A-B}) - (\text{Cov}(\hat{g}^{A-B}, \hat{s}^{A-B}))^2)}{(d_1 - d_2)^2(\text{Var}(\hat{g}^{A-B})\text{Var}(\hat{s}^{A-B}) - (\text{Cov}(\hat{g}^{A-B}, \hat{s}^{A-B}))^2) + \text{Var}(v)} .$$

Also, $p \lim c_1 = \omega \gamma_g + (1 - \omega) \gamma_s$, where the weight ω is determined by the relative covariances of predicted

general and specific human capital with \hat{H}^{A-B} and \hat{h}^{A-B} :

$$\omega = \frac{\text{Var}(\hat{H}^{A-B})\text{Cov}(\hat{g}^{A-B}, \hat{h}^{A-B}) - \text{Cov}(\hat{H}^{A-B}, \hat{g}^{A-B})\text{Cov}(\hat{H}^{A-B}, \hat{h}^{A-B})}{\text{Var}(\hat{H}^{A-B})\text{Var}(\hat{h}^{A-B}) - (\text{Cov}(\hat{H}^{A-B}, \hat{h}^{A-B}))^2}$$

compression and workers bear the full cost and return to general human capital acquisition occurring on the job, so that $b_l = 1$ and $c_l = 0$.

Now subtract (1) from (2) to obtain an equation for wage growth between the first and second period:

$$(5) \quad \ln w_2 - \ln w_1 = (\alpha_2 - \alpha_1) + (\beta_2 - \beta_1) \ln H_1 \\ + (\beta_1 + \gamma_{2g} - \gamma_{1g})g + (\beta_1 + \gamma_{2s} - \gamma_{1s})s + (u_2 - u_1).$$

Consider a regression of $\ln w_2 - \ln w_1$ on human capital acquisition \hat{h} :

$$(6) \quad \ln w_2 - \ln w_1 = b_2 + c_2 \hat{h} + (u_2 - u_1),$$

where $\hat{h} \equiv E(\ln H_2 - \ln H_1 | Z)$. One can show that

$$(7) \quad p \lim c_2 = \beta_1 + (1 - \omega_s)(\gamma_{g2} - \gamma_{g1}) + \omega_s(\gamma_{s2} - \gamma_{s1}) + \omega_2(\beta_2 - \beta_1),$$

where the weight $\omega_s \equiv Cov(\hat{s}, \hat{h})/Var(\hat{h})$ is the contribution of (predicted) specific human capital to the overall variance of (predicted) human capital growth and $\omega_2 \equiv Cov(\hat{H}_2, \hat{h})/Var(\hat{h})$.¹¹

Note that ω_s and ω_2 and therefore c_2 will depend on the choice of instruments.¹² This reflects the fact that the effect of productivity growth on wage growth varies depending on the relative

¹¹ Analogously to above, $\hat{H}_2 = E(\ln H_2 | Z)$. We have :

$$p \lim c_2 = \frac{Cov(\hat{h}, \ln w_2 - \ln w_1)}{Var(\hat{h})} \\ = \beta_1 + \frac{(\beta_2 - \beta_1)Cov(\hat{h}, \ln H_2) + (\beta_1 + \gamma_{g2} - \gamma_{g1})Cov(\hat{h}, \hat{g}) + (\beta_1 + \gamma_{s2} - \gamma_{s1})Cov(\hat{h}, \hat{s})}{Var(\hat{h})}$$

(where we add and subtract $\beta_1 \ln H_2$ from (5) and note that $Cov(u_2 - u_1, \hat{h}) = 0$). The text expression follows by noting that $Cov(\hat{h}, \hat{g}) + Cov(\hat{h}, \hat{s}) = Var(\hat{h})$.

¹² For example, letting $\hat{g} = \kappa_{0g} + \kappa_{1g}Z$, $\hat{s} = \kappa_{0s} + \kappa_{1s}Z$, and noting that

$\Delta \hat{h} = \hat{g} + \hat{s} = (\kappa_{0g} + \kappa_{0s}) + (\kappa_{1g} + \kappa_{1s})Z$, one can easily see that ω_s will be determined by the relative magnitudes of κ_{1g} and κ_{1s} : instruments having a stronger relationship with general human capital relative to specific capital will weight the general human capital coefficient more heavily.

contributions of general and specific human capital to productivity growth.¹³ Under the Becker model, with $\beta_1 = \beta_2 = 1$ and $\gamma_{g2} = \gamma_{g1} = 0$, the coefficient c_2 will converge to $1 + \alpha(\gamma_{s2} - \gamma_{s1})$, so the deviation of c_2 from 1 will depend on the extent to which the employer shares the return to specific capital and (predicted) specific capital's overall contribution to the variance of (predicted) human capital growth. A low value of c_2 can be taken as evidence against the Becker model if it implies an implausibly high contribution of specific capital to overall productivity growth. Note that c_2 can differ from 1 for two possible reasons: (a) because the employer shares the returns to human capital acquisition, either through wage compression or through direct sharing of costs and return (as picked up in the first three terms of (7)) and (b) because the extent of wage compression changes over time.

III. Empirical Results

The wage equations that we estimate are dictated by the workers for whom we have wage and productivity information. In section II we provided interpretations for wage growth regressions and for regressions of differences in starting wages on differences in productivity for workers in the same job. Ideally, we would estimate a wage growth equation for the last worker hired or the typical worker, and we would estimate a separate equation for the difference between the starting wage of the last worker hired and the typical worker. However, we know the typical worker's wage after two years, but not at the start of the job. And while we know the starting wage paid to the last worker hired, we do not have useful information on the wage this worker received subsequently.¹⁴ Consequently, we estimate a pseudo-wage-growth equation, in

¹³ Angrist and Imbens (1996), Angrist, Imbens and Rubin (1995), and other papers in the literature on local average treatment effects analyze the properties of IV when the effect of the regressor of interest varies across the population, and similarly find that the limit of the IV estimator will depend on the choice of instruments.

¹⁴ Employers are asked about the wage received by the last worker hired at the time of the interview, but this worker's tenure is typically quite short. For example, the last worker hired has been on the job less than one year in

which the dependent variable is $\ln(w_{104}^{typ}) - \ln(w_0)$, where w_{104}^{typ} denotes the typical worker's wage after 104 weeks and w_0 denotes the starting wage paid to the last worker hired.

Letting w_0^{typ} denote the typical worker's log starting wage, one can write

$$(8) \quad \ln(w_{104}^{typ}) - \ln(w_0) = (\ln(w_{104}^{typ}) - \ln(w_0^{typ})) + (\ln(w_0^{typ}) - \ln(w_0))$$

The first term on the right hand side of (8) is simply the growth rate in the wage of the typical worker during the first two years of employment. The growth rate in the typical worker's wage depends on the growth in this worker's productivity, or

$$(6') \quad \ln(w_{104}^{typ}) - \ln(w_0^{typ}) = c_2 h^{typ} + l_2 X_2 + \varepsilon_2,$$

where $h^{typ} \equiv \ln(H_{104}^{typ}) - \ln(H_0^{typ})$ denotes the difference in logarithms between the typical worker's productivity after two years and when he is initially hired, and ε_2 is an error term with mean 0. The X_2 term denotes a vector of variables, described below, that influence wage growth independently of productivity, possibly due to contracting considerations. The second term on the right hand side of (8), the difference between the log starting wage of the typical worker and of the last worker hired, depends on these workers' marginal products at the time of hire, or

$$(4') \quad \ln(w_0^{typ}) - \ln(w_0) = b_1 \Delta H + c_1 \Delta h + l_1 X_1 + \varepsilon_1,$$

where $\Delta H \equiv \ln(H_0^{typ}) - \ln(H_0)$ is the difference between the log starting productivity of the typical worker and the last worker hired, $\Delta h \equiv (\ln(H_{104}^{typ}) - \ln(H_0^{typ})) - (\ln(H_{104}) - \ln(H_0))$ is the difference between the productivity growth rate of the typical worker and of the last worker, X_1 denotes variables other than productivity that affect starting wage differences, and ε_1 is a mean zero error term

Substituting (6') and (4') into (8) yields

over 50 percent of the sample observations. We will subsequently use this information to estimate the length of time until the first wage adjustment.

$$(9) \quad \ln(w_{104}^{typ}) - \ln(w_0) = c_2 h^{typ} + b_1 \Delta H + c_1 \Delta h + l_1 X_1 + l_2 X_2 + \varepsilon_1 + \varepsilon_2.$$

OLS Estimation

For comparison purposes, we begin by estimating an OLS version of (9).¹⁵ We do not have a direct report on the productivity of the last worker after two years and therefore are not able to include the variable Δh to obtain an estimate of c_1 . (We include predicted values of Δh in our subsequent analysis.) The control vector X_2 includes a constant, union indicator, the log of the number of employees at the firm, and tenure.¹⁶ Ideally, X_1 would contain differences in characteristics between the typical worker in the job and the last worker hired. However, the survey does not contain information on the characteristics of the typical worker. These characteristics are accordingly imputed by using predicted values from regressions of the characteristics of the last worker hired on the X_2 variables plus a vector Z_{typ} of characteristics of the typical worker or the job that affect productivity but do not otherwise affect wages.¹⁷ Each variable in X_1 is calculated as the difference between the actual characteristic of the last worker hired and the imputed value. The vector X_1 includes the following variables: age, age squared, dummies for years of schooling of 16 or more and greater than 12 but less than 16, and dummies for prior vocational training, female, part-time (less than thirty-five hours a week), and temporary/seasonal job.

As reported in the top panel of Table 2, the estimated coefficient on productivity growth is only 0.035, which, although statistically significant from zero, is very small by any standard. The estimated coefficient on the difference in log productivity, reported in the first column of

¹⁵ Productivity at time 0 equals 0 for a few of the observations. We therefore add a small number (.1) to time 0 productivity and time two year productivity before taking logs. Our results are not sensitive to the choice of number; the results are also not affected if one simply drops the observations where time 0 productivity is zero.

¹⁶ Employers are (implicitly) asked about the starting wage paid to the most recently hired worker *at the time he was hired*. Since wages increase over time, greater tenure is associated with a lower starting wage and higher wage growth.

¹⁷ The vector Z_{typ} is described below.

Table 2, is only 0.023. This is similar to the coefficient on initial productivity of 0.055 found by Bishop (1987) using a sub-sample with observations of two workers in the same job.

The implicit hypothesis in using productivity at the very start of the job as our starting productivity measure is that at any instant workers are paid on the basis of their marginal product at that point in time. This is something of a straw man because wages are not adjusted continuously over time, but instead are fixed for discrete intervals. Indeed, as discussed above, the descriptive statistics reported in Table 1 indicate that productivity growth is particularly high at the very beginning of employment. It seems likely that expected productivity beyond the first two weeks is folded into the starting wage.

The second column of Table 2 reports estimates using average productivity over the first three months on the job as our starting productivity measure, which we calculate as a weighted average of the productivity during the first two weeks and the productivity from the third week to the twelfth week: $H_{0,12} = (1/6)H_{0,2} + (5/6)H_{3,12}$, where $H_{t,t'}$ is average productivity between weeks t and t' . The coefficient on productivity growth increases to 0.073, still quite small, while that on the difference in starting productivity is unchanged.

Instrumental Variables Analysis

We now turn to instrumental variable (IV) estimation. IV has two advantages here: it corrects for measurement error in productivity, a likely reason for the small estimated coefficients using OLS, and it gives us a measure for predicted Δh . to allow estimation of (9).

As noted by Bishop (1987), the EOPP productivity estimates likely contain considerable measurement error. Prima facie evidence for this is provided by the fact that the productivity reports are predominantly round numbers. For example, about 95% of the 1,502 observations for the productivity of the last worker hired in the first two weeks are divisible by 5, and about 5%

of the observations are exactly 50. Measurement error typically implies that OLS estimates are biased downward. However, one might well argue that the standard result does not apply to the present case because the wage offered by an employer should depend directly on the employer's imperfect estimate of worker productivity, and not on the worker's unobserved "true" productivity.

There are two possible explanations why the employer's wage offer might depend on something other than the employer's reported productivity estimate. First, the employer might recognize that his initial estimate of the worker's productivity, H , is fraught with error, and refine it by combining it with the average productivity of workers with similar characteristics to obtain an improved estimate, \tilde{H} , on which pay is based. To the extent that the employer reports H rather than \tilde{H} to the EOPP survey, the estimated OLS effect of productivity on wages will be biased downward.

The alternative explanation is "Friedmanesque" in nature: employers may simply be acting as if they are making wage offers based on their assessments of worker productivity. Rather than consciously forming productivity estimates, employers may be basing their wage offers on workers' observed characteristics, and presumably being driven to the optimal wage offers over time. In this scenario, employers only explicitly obtain an actual measure of worker productivity when they are asked to do so by the survey. Again, the estimated OLS effect of productivity on wages will be biased downward.

We adopt a two stage least squares procedure. In the first stage, we estimate a productivity equation, which is then subsequently used to predict h^{yp} , ΔH , and Δh in (9). Our productivity equation estimates the path that productivity follows over time. Thus, besides eliminating the possible downward bias from measurement error, our IV analysis also enables us

to examine the effect of allowing the length of time until the first wage adjustment--hereafter referred to as "contract length"--to vary. For the time being, we continue to assume that the wage is initially set for a period of twelve weeks.

Employers in the survey are asked about the productivity of the last worker hired at the time of the interview and the typical worker who has been in the position for two years. This allows us to estimate the following productivity equation

$$(10) \quad \ln(p_t / p_{104}^{typ}) = \psi X_1 + \delta X_2 + \chi_{typ} Z_{typ} + \chi Z + (\psi_1 X_1 + \chi_{1typ} Z_{typ} + \chi_1 Z') \tau_1 \\ + (\psi_2 X_1 + \chi_{2typ} Z_{typ} + \chi_2 Z') \tau_2 + m + v_t,$$

where we adopt the spline specification $\tau_1 = \min(\ln(t+1), \ln(12+1))$ and $\tau_2 = \max(\ln(t+1) - \ln(12+1), 0)$, tenure t is measured in weeks, and where we have normalized by dividing the last worker's productivity by the productivity of the typical worker after two years. The vector Z_{typ} includes the logarithm of the number of hours the typical new employee in the position spends watching other people do the job, the logarithm of the number of hours of formal training provided to the typical worker, the logarithm of the number of hours of informal training provided to the typical worker, and the logarithm of the number of weeks it takes a new employee in the most recently filled position to become fully qualified. The vector Z contains the log of the *most recently hired* worker's formal training, the log of the *most recently hired* worker's informal training, and the log of the *most recently hired* worker's relevant experience; the vector Z' excludes the log of relevant experience from Z . The residual consists of m , which is a mean-zero person-specific fixed effect, and v_t , which is iid measurement error. The error structure is important in allowing us to derive a closed-form expression for expected log productivity growth, as the fixed effect

factors out.¹⁸ This error structure implies that the variance of the regression error is invariant with respect to tenure, as is the covariance between residuals for the same worker at different tenures. Neither implication is rejected by the data.¹⁹

From (10), the most recently hired worker's true productivity during the first twelve weeks is equal to

$$\begin{aligned}
 (11) \quad p_{0,12} &\equiv \int_0^{12} \frac{(p_t - v_t)}{12} dt \\
 &= p_{104}^{typ} \exp(m + \psi X_1 + \delta X_2 + \chi_{typ} Z_{typ} + \chi Z) \int_0^{12} \frac{\exp(\psi_1 X_1 + \chi_{1typ} Z_{1typ} + \chi_1 Z') \ln t}{12} dt \\
 &= p_{104}^{typ} \exp(m) J(Z_{typ}, Z, Z', 12, 0),
 \end{aligned}$$

$$\text{where } J(y_1, y_2, y_3, y_4, y_5) \equiv \frac{\exp(\psi X_1 + \delta X_2 + \chi_{typ} y_1 + \chi y_2) ((y_4 + 1)^{\psi_1 X_1 + \chi_{1typ} y_1 + \chi_1 y_3} - (y_5 + 1))}{(y_4 - y_5) (\psi_1 X_1 + \chi_{1typ} y_1 + \chi_1 y_3 + 1)}.$$

Note that the productivity of the typical worker during the first 12 weeks can be obtained by substituting the typical worker's training and experience for that of the last worker hired.

Specifically,

$$(12) \quad p_{0,12}^{typ} = p_{104}^{typ} \exp(m) J(Z_{typ}, Z_{typ}^n, Z_{typ}', 12, 0),$$

¹⁸ One can show that interacting imputed variables with the tenure spline leads to an error structure that is heteroscedastic with respect to tenure so that there is no closed form solution for predicted productivity growth (equation (13) below). This is why X_2 and the log of relevant experience are not interacted with the tenure splines. The p value of the interaction of these variables with the tenure spline variables is .25.

¹⁹ The variance of the regression error in starting productivity during the first two weeks is significantly larger than the variance of the regression error for all subsequent points in time, but the first two weeks constitute a sufficiently small fraction of the first period that the fixed-effect is still a good approximation. Regressing the squared residual on tenure and tenure squared beyond the first two weeks yields a p value of .59. Interacting the residual for starting productivity with the residual for productivity from the third to the twelfth week and regressing on tenure and tenure squared yields a p value of .39; interacting the residual for starting productivity with the residual for productivity after the twelfth week and regressing on tenure and tenure squared yields a p value of .93; interacting the residual for the residual for productivity from the third to the twelfth week with the residual for productivity after the twelfth week and regressing on tenure and tenure squared yields a p value of .83.

where Z'_{typ} denotes the vector obtained when one replaces the elements of Z' with the corresponding elements of Z_{typ} and Z''_{typ} denotes the vector obtained by replacing the elements of Z with the corresponding elements of Z_{typ} and the relevant experience of the typical worker. From (12), the predicted log productivity growth of the typical worker is given by

$$(13) \quad E(\ln(p_{104}^{typ}) - \ln(p_{0,12}^{typ}) | X_1, X_2, Z_{typ}, Z'_{typ}, Z''_{typ}) = -\ln(J(Z_{typ}, Z''_{typ}, Z'_{typ}, 12, 0)).$$

One can similarly predict $\ln(p_{0,12}^{typ}) - \ln(p_{0,12})$, the difference between the log productivity of the most recently hired worker and the typical worker during the first twelve weeks, and

$$(\ln(p_{104}^{typ}) - \ln(p_{0,12}^{typ})) - (\ln(p_{104}) - \ln(p_{0,12}^t)), \text{ the difference in log productivity growth}$$

Selected coefficients in the estimated productivity equation can be found in Table 3.²⁰

The coefficients conform to one's expectations. The greater the last hire's relevant experience in jobs having some application to his current position, the greater is the initial productivity of the last hire relative to the productivity of the typical worker who has been in the job two years. In contrast, the length of time it takes the typical worker with no experience to become fully trained and qualified, the number of hours the typical worker spends in training activities watching other people do the job, and the training the last worker hired receives all reduce the initial productivity of the last hire relative to the productivity of the typical worker who has been in the job two years. These variables also raise the rate at which the relative productivity of the last worker hired rises over time.²¹ The latter effect is much stronger during the first three months than subsequently.

²⁰ Note that we have up to three observations per worker--an observation from between 0 and 2 weeks, from 3 to 12 weeks, and productivity at the time of the interview.

²¹ Relevant experience and the length of time until qualified are ideal instruments in that they should only affect wages indirectly through their effect on productivity. However, aside from its effect on productivity, training may entail costs that also affect wages. We would expect this effect to be relatively small compared to the productivity effect. But in any case, given the positive correlation between training and productivity growth and the negative

Table 4 shows IV estimates of the wage equation. As indicated in the first row of Table 4, the IV estimate of b_1 is equal to .32. While significantly different from zero, it is also substantially and significantly below one. Recall that the difference between one and this coefficient indicates the degree of wage compression for starting wages, so we have evidence of substantial wage compression. The estimate of c_1 is only equal to .19. Recall that c_1 indicates the degree of employer sharing of the costs of general human capital acquisition, so there is little evidence of sharing. However, this coefficient is imprecisely estimated, so one cannot rule out substantial sharing. The IV estimate of c_2 is 0.22, indicating that wage growth is substantially less than productivity growth.

The preceding sharing rule estimates assume that productivity over the first twelve weeks is folded into the starting wage. It is straightforward to modify the analysis to reflect alternative assumptions about contract length.²² Rows 2- 5 of Table 4 show estimates of the productivity coefficients for other values of τ . The b_1 coefficient is quite stable, showing that the evidence for wage compression is not an artifact of a particular assumed contract length.. The c_1 and c_2 coefficients are more volatile, with c_1 decreasing (and even becoming negative) and c_2 increasing with contract length.

One obvious objection to the previous estimates is that contract lengths need not be constant across workers. Fortunately, one can estimate contract length from the data in the

correlation between training and starting productivity, this will cause one to overestimate both the effect of productivity growth on wage growth and the effect of starting productivity on starting wages.

²² If wages are initially set for a period of length $\tau \geq 12$, then starting productivity over the initial contact period is

given by $p_{0,\tau} = \frac{12}{\tau} p_{0,12} + \frac{(\tau - 12)}{\tau} p_{\tau,12}$, so that the predicted productivity growth of the typical worker is

given by

$$\ln(p_{104}^{typ}) - \ln(p_{0,\tau}^{typ}) = -\ln\left(\left(\frac{12}{\tau}\right)J(Z_{typ}, Z_{typ}''', Z'_{typ}, 12, 0) + \left(\frac{\tau - 12}{\tau}\right)J(Z_{typ}, Z_{typ}''', Z'_{typ}, \tau, 12)\right).$$

Similarly, one can obtain expressions for the predicted difference between the log productivity of the most recently hired worker and the typical worker during the first τ weeks and for the predicted difference in log productivity growth.

survey since EOPP provides information on the last-hired worker's wage at the time of the interview. We assume that the probability that a worker with a specified amount of tenure, t , has already experienced a wage change is given by

$$(14) \quad \Pr(w_t \neq w_0) = F(\xi_t),$$

where the function $F(\cdot)$ is the cumulative standard normal distribution and where the vector ξ_t includes the variables tenure, tenure squared, a dummy variable equal to 1 if tenure exceeds one year, predicted productivity growth after the initial twelve week period ($E(\ln(p_{104}^{yp}) - \ln(p_{0,12}) | X_1, X_2, Z_{yp}, Z'_{typ}, Z''_{typ}))$), and tenure interacted with predicted productivity growth. Productivity growth is included in the probit wage change equation because it seems likely that the faster productivity grows, the shorter will be the length of the period for which the wage is fixed.²³

All of the productivity variables in (9) are functions of the contract length τ . Integrating across possible contract lengths using the parameter estimates in (10) and (14), a worker's expected wage growth is given by

$$(15) \quad \begin{aligned} E(\ln(w_{104}^{yp}) - \ln(w_0)) &= c_2 (\ln(\hat{H}_{104}^{yp}) - \int_0^{\infty} \ln(\hat{H}_0^{yp}(\xi_t)) f(\xi_t) dt) \\ &+ b_1 \int_0^{\infty} ((\ln(\hat{H}_0^{yp}(\xi_t)) - \ln(\hat{H}_0(\xi_t))) f(\xi_t) dt) \\ &+ c_1 \int_0^{\infty} \Delta \hat{h}(\xi_t) f(\xi_t) dt + l_1 X_1 + l_2 X_2. \end{aligned}$$

In practice, we assume that the initial wage is not set for a period longer than 104 weeks or shorter than 12 weeks and amass the probability weights in the tails at the truncation points.²⁴

²³ Our estimation results are not sensitive to the choice of τ .

²⁴ The estimated probit parameters imply that on average the probability that the wage is initially set for a period of time that is less than or equal to 12 weeks is .41. Since the productivity growth coefficient is positively related to

The probit estimation results are reported in Table 5 and the resulting estimates of b_1 , c_1 , and c_2 are reported in the last row of Table 4. The probit results indicate that at least in this data set, wages are initially set for a relatively short period of time: for the average worker, mean contract length is only 18 weeks.²⁵ The wage change probit also provides support for the hypothesis that the probability of wage changes varies with expected productivity--a joint test rejects the hypothesis of no effect of the productivity coefficients at the 1 percent level. For employees with the mean level of productivity growth, the predicted probability of a wage change by 18 weeks is 50 percent, while for employees with productivity growth one standard deviation above the mean the corresponding probability is 55 percent.

As reported in the last row of Table 4, the resulting estimate of b_1 is .32 and the estimate of c_2 is .26. Both estimates are significantly different from one. The estimate of c_1 is only .17, but is very imprecise. Thus, the data provide strong evidence of wage compression, but do not allow us to say much about the degree of cost sharing in the form of a positive relationship between the starting wage and productivity growth. Interestingly, when productivity growth is zero, the estimated two year wage growth of the average worker in the average job is 6 percent. This may reflect the fact that the employer's productivity estimate does not incorporate all costs

tenure for low values of tenure, truncating the contract length distribution at 12 weeks raises the estimated coefficient on productivity growth. The estimated probability that the wage is initially set for a period in excess of two years is .13. The results in Table 4 imply that truncating the contract length distribution at 104 weeks has little effect on the estimated productivity growth coefficient.

²⁵ One possible objection to this analysis is that the data are from the early 1980s, at the end of a period of high inflation. Many of the raises for low values of tenure may have been cost-of-living adjustments unrelated to productivity. We tested for this by estimating a non-linear regression of real wage changes from the start of the job to the interview date on productivity changes weighted by the probability of a true, non-cost-of-living raise (estimated as part of the regression). Our point estimates indicate that all raises were true, productivity-related raises. Specifically, we assumed a wage equation of the form

$$\ln(w_t) - \ln(w_0) = k + \int_{12}^{\tau} \varphi(\mu, \sigma, \tau) c(\ln(p_{\tau,52}) - \ln(p_{0,\tau})) d\tau + \eta,$$

indicating the probability of a true wage increase at τ , $\ln(p_{\tau,52})$ is the log of average productivity from time τ until 52 weeks, and $\ln(p_{0,\tau})$ is the log of average productivity from time 0 until τ weeks. Our point estimates of μ and σ are 13 and .7 if we use the entire sample and 13 and .9 if we restrict the sample to those who have a positive wage change), again indicating that that contract length is typically quite short.

associated with searching for, hiring, and training new workers.²⁶ As in Salop and Salop (1976), these costs may be factored into the starting wage (instead of being spread out into future wages).

What is the source of identification for the b_1 coefficient? To answer this question, we regress $\Delta H \equiv \int_0^{\infty} ((\ln(\hat{H}_0^{byp}(\xi_t)) - \ln(\hat{H}_0(\xi_t)))f(\xi_t)d\xi_t)$ on the other variables in (9) to generate residuals capturing the identifying variation in ΔH . Regressing these residuals on the omitted instruments other than the log of relevant experience yields an R^2 of .068. Adding the log of relevant experience to this regression raises the R^2 to .793, indicating that the identifying variation in predicted ΔH is predominantly variation in relevant experience.²⁷

Recent theoretical work by Acemoglu and Pischke (1999) and Booth and Zoega (2004) has suggested that wage compression may make training more likely, arguing that employers are more willing to finance training if the returns to training accrue at least partially to the employer due to wages only partly responding to increases in productivity. Empirical analysis of this effect has been hampered by the lack information on productivity.²⁸ To investigate this issue we divided the sample by hours of training for the typical worker and reestimated the variable contract length model. The estimates of b_1 show little difference between the subsamples. Estimated c_2 is 0.13 for the sample with training above the sample median and 0.32 for the low-training subsample, but the estimates are not significantly different from each other ($t=1.22$).

²⁶ One issue with interpreting this result is that it could at least partially be an artifact of functional form, as we assume a linear relationship between log productivity growth and log wage growth and only a few observations have zero or near-zero productivity growth. Further investigation of this question is outside the scope of this paper.

²⁷ The remaining unexplained variation is accounted for by the non-linear construction of the variable.

²⁸ Empirical studies include Brunello (2002), Almeida-Santos and Mumford (2004), Ericson (2004), and Peraita (2001). The findings are mixed.

IV. A Simple Sharing Model with Equity Norms

The foregoing results imply that predictable differences in employees' starting productivity based on their relevant job experience are not completely reflected in wages. This is contrary to traditional theory, which suggests that any rents earned by firms by paying experienced hires less than their productivity would be bid away by competing employers.

While one can find explanations of wage compression in the literature, these seem for the most part unsatisfactory in the present context. Frank (1984) proposes a model where more productive workers derive utility from their status within the firm and are willing to accept a wage less than their marginal product in order to work side by side with workers who are paid less. However, it seems unlikely that workers who are more productive simply because they are more experienced would derive utility from this fact. Similarly, Akerlof and Yellen (1990) posit that large wage differentials between groups may be perceived as unfair and lead to reduced effort. However, since less-experienced workers will acquire experience (barring layoffs) it is not clear why they would feel aggrieved by steep wage-experience profiles. Explanations based on asymmetric information, as in Acemoglu and Pischke (1998,1999), Katz and Ziderman (1999), and Chang and Wang (1996) fail to account for the fact that predicted productivity as a function of relevant experience--a variable which is publicly observable--is not rewarded proportionately.

As discussed by Acemoglu and Pischke (1999) and Zoega and Booth (2005), monopsony power by employers can lead to wage compression if search is more costly for higher skilled workers. However, this seems unlikely by itself to explain the magnitude of starting wage compression that we observe in the data. In fact, if one allows entry, the monopsony model does not unambiguously predict compression in the starting wage. There are two competing effects.

A more highly skilled worker for whom search is more costly will be in a weaker bargaining position than a less skilled worker with lower search costs. But knowing that he will be relatively immobile once he accepts the employer's job and thus subject to higher rent extraction by the employer in the future, the more highly skilled worker will bargain for a higher proportion of his compensation up front. If the second effect dominates the first, starting wages will vary *more than proportionately* with productivity.

Here we suggest an explanation based on the interaction of equity considerations such as those in Akerlof and Yellen (1990) with employer sharing of the cost of human capital acquisition. If employers share the costs of human capital investment, they recoup these costs by paying employees trained by the firm less than their marginal product. Hiring trained outsiders at a wage equal to their marginal product implies that observationally equivalent workers are paid different wages. It is easy to imagine that an employer's senior workers will be unhappy and put forth less effort if they receive a lower wage than other experienced workers who *are no more productive*, but who simply began their careers at other firms. Such behavior seems consistent with both casual observation of the labor market and experimental studies cited in Akerlof and Yellen (1990).²⁹ We show that an "equity norm" will not only lead to wage compression, but also will amplify employer sharing of the return to general human capital acquisition, as the wage compression and sharing effects will reinforce each other.

Our model is similar to that developed by Hashimoto (1981) to analyze the optimal division of the returns and costs to firm-specific human capital investment. Unlike Hashimoto (1981), we assume that human capital is purely general in that its expected value is the same

²⁹ In their experimental study, Charness and Kuhn (2005) find little response of effort to co-worker wages, but explicitly do not consider the case where workers know that they are identically productive. In our analysis, this would not be a "knife-edge" case since employers paying senior workers less than newly hired experienced workers would be differentiating on the basis of a characteristic that is known to be unrelated to productivity.

everywhere. Consider a match between an employer and an employee who is in the labor market for two periods. The worker is hired in period 1 and has a starting value of marginal product of H . As a result of on-the-job training and learning by doing, the worker's expected productivity is higher in period 2. The worker's actual period 2 productivity is random and is not observed until the start of that period. Letting h denote the value in period 2 of the human capital accumulated in the initial period and η denote the value of the productivity shock, the worker's actual value of marginal product in period 2 is $H + h + \eta$.

The worker's utility in each period is equal to the wage he receives plus the amenity value ε he places on the employer's job. The worker observes ε after he starts the job. The worker receives a starting wage equal to w_1 and a second period wage of w_2 . If the worker moves to a new job, he incurs the moving cost $c = c(H + h)$. Assuming that the worker's human capital is general, he receives the alternative wage $H + h$.

We assume that the cost of locating and moving to a new job increases with the worker's stock of human capital. More senior positions at an employer are typically filled from within by workers who have proved to be a good match. Consequently, a worker with more experience and training searching for a job typically has a smaller set of openings available to him and will generally find it more difficult and costly to find and relocate to an employer who can use his skills effectively.³⁰

³⁰ This motivation differs from that in Acemoglu and Pischke (1999), who argue that locating another job requires a period of unemployment, which is more costly to higher paid workers. Zoega and Booth's (2005) model of wage compression assumes that the set of employers who can use more able, higher skilled workers is smaller than the set of employers who can use less skilled workers. This is more similar to our motivation except that Zoega and Booth's argument is applicable when one is comparing workers in high- and low- skilled occupations. In contrast, we are comparing workers who are in the same basic occupation.

Because the value of η is the private information of the employer and the value of ε is the private information of the worker, we assume, like Hashimoto, that the two parties cannot contract on the basis of the realized values of η and ε .³¹ Letting

$$(16) \quad D = H + h - w_2$$

denote the rent the employer extracts when η is zero, the employer dismisses the worker if $\eta < -D$. It pays for the worker to quit when $\varepsilon < \varepsilon^c = D - c$. The probability of a dismissal or layoff

is $L = \int_{-\infty}^{-D} g(\eta) d\eta$ and the probability of a quit are thus given by $Q = \int_{-\infty}^{\varepsilon^c} f(\varepsilon) d\varepsilon$, where $g(\cdot)$ and $f(\cdot)$

denote the density functions of η and ε .

The expected gain to the worker from his match with the employer is given by

$U = w_1 + \delta((1-L)(1-Q)E(w_2 + \varepsilon | \varepsilon \geq \varepsilon^c) + ((1-L)Q + L)(H + h - c))$, where δ denotes the discount factor. The worker is willing to form a match with the employer if U is at least as great as the expected utility U^A available to experienced workers elsewhere in the labor market. The employer's expected gain from his match with the worker is $\pi = H - w_1 + \delta(1-L)(1-Q)E(D + \eta | \eta \geq -D)$.

The employer chooses first and second period wages to maximize π subject to the constraint that $U \geq U^A$. Note that the first period wage simply serves to divide up the total return to the match between the employer and worker. In a competitive labor market, the first period wage is bid up until the employer's expected profit over the two periods is driven to zero, or

$$(17) \quad w_1 = H + \delta(1-L)(1-Q) E(D + \eta | \eta \geq -D).$$

³¹ Mortensen (1978) shows that inefficient separations can be eliminated if the employee and employer post turnover bonds that provide compensation for losses imposed by the other's decision to separate. While one occasionally observes turnover bonds, as noted by Black and Loewenstein (1998), their general use is limited by the fact that they are hard to implement when the exact value of the match to one party is not known by the other party. In addition, as noted by Carmichael (1983), turnover bonds suffer from the disadvantage of providing the worker and employer with an incentive to induce the other party to initiate turnover.

When the employer decides whether or not to dismiss a worker, he does not take into account the potential loss that the dismissal imposes on the worker. Similarly, when the worker decides whether to quit, he does not take into account the potential loss that a quit imposes on the employer. The optimal second period wage, or, equivalently, the optimal D , minimizes the expected loss from inefficient separations. Differentiating and rearranging terms one finds that at the optimum

$$(18) \quad f(\varepsilon^c) \int_{-D}^{\infty} (\eta + D)g(\eta)d\eta = g(-D) \int_{\varepsilon^c}^{\infty} (\varepsilon - \varepsilon^c)f(\varepsilon)d\varepsilon,$$

or equivalently, $(\partial Q/\partial D)(1-L)E(\eta+D|\eta>-D) = -(\partial L/\partial D)(1-Q)E(\varepsilon-\varepsilon^c|\varepsilon>\varepsilon^c)$: the reduction in the worker's expected capital loss from a marginal increase in D must just equal the increase in the employer's expected capital loss.³²

Let D^0 denote the optimal value of D . Differentiating (17) with respect to h , one finds that

$$(19) \quad \partial D^0 / \partial h = Mc'(H + h),$$

where $0 < M < 1$.³³ So if greater human capital accumulation is associated with higher mobility costs, the employer will share part of the return with the worker. Differentiating (17) yields

$$(20) \quad \partial w_1 / \partial h = \delta c'((1-Q)M + f(\varepsilon^c)(1-M) \left(\int_{-D^0}^{\infty} (D^0 + \eta g(\eta))d\eta \right) > 0:$$

³² The second-order condition to the maximization condition requires that $\partial^2 G / \partial D^2 < 0$, which is equivalent to

$A + B > 0$, where $A \equiv g(-D) \int_{\varepsilon^c}^{\infty} f(\varepsilon)d\varepsilon + f'(\varepsilon^c) \int_{-D^c}^{\infty} (D + \eta)g(\eta)d\eta$ and

$B \equiv f(\varepsilon^c) \int_{\eta^c}^{\infty} g(\eta)d\eta + g'(-D) \int_{\varepsilon^c}^{\infty} (\varepsilon - \varepsilon^c)f(\varepsilon)d\varepsilon$. This condition is satisfied provided that the responsiveness

of the quit rate and of the dismissal rate does not decrease too much as the wage rate increases. Henceforth we assume that A and B are both positive.

³³ Specifically, $M = \frac{A}{A + B}$.

the higher rent earned by the employer in the second period is paid to the worker in the form of a higher first period wage (generating a positive value for the coefficient λ_g in equation (1) above, unlike the Becker model). Finally, differentiating (17) with respect to H yields

$$(21) \quad \partial w_1 / \partial H = 1 + \delta(1 + \phi')c'(M(1 - Q) + f(\varepsilon^c)(1 - M)) \left(\int_{-D^0}^{\infty} (D^0 + \eta g(\eta)) d\eta \right) > 1.$$

The increase in mobility costs from the higher H leads to rent extraction in the second period, which is compensated for by a higher wage in the first period. Thus the variation in the starting wage *exceeds* the variation in starting productivity, contrary to our empirical results. Note that we have assumed that an employer hiring a new worker does not have any monopsony power, so that the newly hired worker's expected lifetime compensation equals his expected lifetime value of marginal product. As mentioned above, allowing monopsony power breaks this equality, with the result that there is no clear prediction as to whether variation in the starting wage exceeds or is less than variation in starting productivity.

Equity Norms

We now posit the existence of an "equity norm" that prevents an employer from paying retained workers less than equally productive experienced workers hired from the outside. Such a norm implies that a worker who moves to a new employer in period 2 will realize a smaller return to his initial general human capital investment and will be less likely to quit. Specifically, letting the wage the worker can receive at an alternative employer be given by $H + h - D^A$, the minimum value of ε such that the worker does not quit is now given by $\varepsilon^c = D - D^A - c$.

Other than the changed definition ε^c , condition (18) is unchanged. This condition implicitly defines D as a function of D^A : $D = \chi(D^A)$. Differentiating D with respect to D^A , one

finds that $\chi' = M$: a reduction in the alternative wage causes a partial reduction in the second period wage offered by a worker's initial employer.

Competition for experienced workers will ensure that they do not receive less than retained workers. Thus, if all employers hire a mix of inexperienced and experienced workers, labor market equilibrium requires that $D = D^A$. Note that when $D = D^A$, $\varepsilon^c = -c$. The equity norm equilibrium conditions therefore reduce to

$$(22) \quad f(-c) \int_{-D^c}^{\infty} (D + \eta)g(\eta)d\eta = g(-D) \int_{-c}^{\infty} (c + \varepsilon)f(\varepsilon)d\varepsilon .$$

Let D^* denote the equilibrium value of D . It follows immediately that $D^* < D^0$.

Unlike the case without equity constraints, and consistent with our empirical results, the equity norm implies that the variation in starting productivity will generally exceed the variation in the starting wage. For example, suppose that starting differences in the stock of human capital are fully offset by differences in human capital accumulation during the first period of employment and let \bar{H} denote the stock of human capital achieved by all workers in the second period. Then a new worker will not receive compensation for any starting human capital that exceeds $\bar{H} - D^*$.

The existence of an equity norm also affects the extent to which employers share in the return to human capital accumulation. Specifically, differentiating (22), one finds after a little algebra that

$$(23) \quad \partial D^* / \partial h = Mc'(h)/(1 - M) .$$

Comparing (23) and (19), one sees that the existence of an equity norm amplifies the initial tendency by employers to share in the costs and returns to general human capital investment. In

terms of our empirical model, this implies a reduction in c_2 and an increase in c_1 . While c_1 is not precisely estimated, our model is consistent with low values of b_1 and c_2 .

In deriving the equity-norm equilibrium, we have assumed that employers hire a mix of inexperienced and experienced workers. This will be the case if inexperienced and experienced workers are strong complements in production, something which seems consistent with casual observation. For example, it may be efficient to place less skilled, inexperienced workers in less demanding tasks and let experienced workers concentrate on certain critical tasks for which they are better suited. Or part of experienced workers' higher value may stem from the informal training that they provide to less experienced workers.³⁴

V. Conclusion

Exploiting the richness of the EOPP data, this paper has analyzed the relationship between wages and productivity during the early years of an employment relationship. The EOPP data show that worker productivity grows substantially during the early part of the employment relationship: productivity after two years is on average eighty percent higher than at the start of the job. The data also indicate that most of the growth in productivity occurs at the very start of the job, with sixty-four percent of the growth in two year productivity occurring during the first three months.

Even when one corrects for measurement error and the fact that expected productivity beyond the start of the job may be folded into the starting wage if wage revisions are not instantaneous, one finds that variation in productivity is only partially reflected in wages. One observes this along two dimensions. First, there is compression in the starting wage. Starting productivity differences for workers in the same job – in large part driven by differences in

³⁴ In their formally similar model, Akerlof and Yellen (1990) consider equilibria with segregation of different types of workers; detailed analysis of these equilibria in the present case is outside the scope of this paper.

relevant experience - are only partially reflected in starting wage differences even when one controls for differences in human capital acquisition. Second, productivity growth stemming from human capital accumulation while on the job is only partially reflected in wage growth, reflecting employer sharing of the costs and returns to human capital acquisitions and/or increased wage compression over time.

Our empirical findings can be explained by adding equity norms to a simple model of employer sharing of the costs of human capital. Our model has two key features. First, similar to Zoega and Booth (2005), the cost of locating and moving to a new job increases with the worker's stock of human capital, reflecting the fact that more highly skilled workers typically have a smaller set of potential employers to choose from. Second (and to our knowledge unique to our model), in offering wages to new workers, an employer is constrained by the fact that his senior workers will be unhappy and uncooperative if they receive a lower wage than other workers who are no more productive. In equilibrium, there can be substantial wage compression: higher productivity resulting from the acquisition of human capital is only partially reflected in wages, whether the human capital accumulation occurs on the current job or at other jobs.

Similar to Acemoglu and Pischke (1999) and Zoega and Booth (2005), a key feature of our model is the association of wage compression with employers' sharing the cost of human capital acquisition. Our estimate of the sharing effect is right signed, but the data are unfortunately insufficient to estimate the sharing effect precisely. Acemoglu and Pischke (1999) and Booth and Zoega (2004) suggest that compression may be associated with willingness of firms to train workers. Unfortunately, the data do not allow precise estimation of the difference in wage compression between low- and high-training firms. Thus, while the presence of wage compression receives strong support in our data, we have little to say empirically about the

implications of this finding. Further empirical analysis is required to test models associating wage compression with employer cost sharing or employer provision of training. This analysis is made difficult by the paucity of data sets with information on productivity.

References

- Acemoglu, Daron and Jorn-Steffen Pischke. 1998. "Why do Firms Train: Theory and Evidence." *The Quarterly Journal of Economics* Vol. 113, No. 3 (February), pp. 79-119.
- Acemoglu, Daron and Jorn-Steffen Pischke. 1999. "The Structure of Wages and Investment in General Training." *Journal of Political Economy* Vol. 107, No. 3 (June), pp. 539-572.
- Akerlof, George A. and Janet Yellen. 1990. "The Fair-Wage Effort Hypothesis and Unemployment." *The Quarterly Journal of Economics* Vol. 102, No. 2 (May), pp. 255-283.
- Almeida-Santos, Felipe and Karen Mumford. 2004. Employee Training and Wage Compression in Britain," IZA Discussion Paper No. 1197.
- Angrist, Joshua D. and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* Vol. 90, No. 430. (June), pp. 431-442.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* Vol. 91, No. 434 (June), pp. 444-455.
- Autor, David. 2001. "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics* Vol. 116, No. 4 (Nov.), 1409-1448.
- Balmaceda, Felipe. 2005. "Firm-Sponsored General Training," *Journal of Labor Economics* Vol. 23 No. 1 (January), pp. 115-134.
- Barron, John M., Mark C. Berger, and Dan A. Black. 1999. "Do Workers Pay for On-the-Job Training?" *Journal of Human Resources* Vol. 34, No. 2, pp. 235-252.
- Bartel, Ann P. 1995. "Training, Wage Growth, and Job Performance: Evidence from a Company Database." *Journal of Labor Economics* Vol 13, No.3 (July), pp. 401-425.
- Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* Vol. 70 Supplement (October), pp. 9-49.
- Bishop, John. 1987. "The Recognition and Reward of Employee Performance." *Journal of Labor Economics* Vol. 4, No. 4, Part 2 (October), pp. S36- S56.
- Bishop, John. 1991. "On-the-Job Training of New Hires" in *Market Failure in Training*, edited by David Stern and Josef Ritzen, New York, Springer Verlag, pp. 61-96
- Black, Dan A. and Mark A. Loewenstein. 1997. "Dismissals and Match-Specific Rents." *Labour Economics: An International Journal* Vol 4, No. 4 (December), pp. 325-340.

- Bollinger, Christopher R., and Amitabh Chandra. 2005. "Iatrogenic Specification Error: A Cautionary Tale of Cleaning Data." *Journal of Labor Economics* Vol. 23. No. 2 (April), pp. 235-257.
- Booth, Allison L. and Mark L. Bryan. 2002. "Who Pays for Training: New Evidence for British Men and Women," IZA Discussion Paper No. 486.
- Booth, Allison L. and Gylfi Zoega. 2004. "Is Wage Compression a Necessary Condition for Firm-Financed General Training?" *Oxford Economic Papers* Vol. 56, pp. 88-97.
- Brunello, Giorgio, "Is Training More Frequent when Wage Compression is Higher? Evidence from 11 European countries," The Fondazione Eni Enrico Mattei Note di Lavoro Series 10.2002.
- Carmichael, H. Lorne. 1983. "Firm-Specific Human Capital and Promotion Ladders." *Bell Journal of Economics* Vol. 14, No. 2 (Spring), pp. 251-58.
- Chang, Chun and Yijiang Wang. 1996. "Human Capital Investment Under Asymmetric Information: The Pigovian Conjecture Revisited." *Journal of Labor Economics* Vol. 14 No. 3 (July), pp. 505-519.
- Charness, Gary and Peter Kuhn, 2005. "Pay Inequality, Pay Secrecy, and Effort: Theory and Evidence," Working Paper, University of California, Santa Barbara.
- Ericson, Thomas. 2004. "The Effects of Wage Compression on Training: Swedish Empirical Evidence," Institute for Labour Market Policy Evaluation Working Paper No. 2004.15.
- Frank, Robert H. 1984. "Are Workers Paid their Marginal Product?" *American Economic Review* Vol. 74, No. 4 (September), pp. 549-571.
- Frazis, Harley and Mark A. Loewenstein. 2005. "Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation" *Journal of Human Resources* Vol. 40, No. 2 (Spring), pp 453-476.
- Harhoff, D. and T.J. Kane. 1997. "Is the German Apprenticeship System a Panacea for U.S. Labor Market?" *Journal of Population Economics* Vol 10, pp. 171-196.
- Hashimoto, Masanori. 1981. "Firm-Specific Investment as a Shared Investment." *American Economic Review* vol. 71, No. 3 (June), pp. 475-482.
- Katz, Eliakim and Adrian Ziderman. 1990. "Investment in General Training: The Role of Information and Labour Mobility." *The Economic Journal*, pp. 1147-58.
- Loewenstein, Mark A. and James R. Spletzer. 1998. "Dividing the Costs and Returns to General Training," *Journal of Labor Economics* Vol. 16 No. 1 (January), pp. 142-171.

- Loewenstein, Mark A. and James R. Spletzer. 1999. "General and Specific Training: Evidence and Implications," *Journal of Human Resources* Vol. 34, No. 4 (Fall), pp. 710-733
- Medoff, James and Katherine Abraham. 1981. "Are Those Paid More Really More Productive? The Case of Experience." *Journal of Human Resources* Vol. 16, pp. 186-216.
- Mortensen, Dale T. 1978. "Specific Capital and Labor Turnover" *Bell Journal of Economics* Vol. 9, No. 2 (Autumn), pp. 572-86.
- Peraita, Carlos. 2001. "Testing the Acemoglu-Pischke model in Spain," *Economics Letters* Vol. 72, No. 1, pp. 107-115.
- Salop, Joanne and Steven Salop. 1976. "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics*, Vol. 90, No. 4 (Nov., 1976), pp. 619-627
- Schmidt, Frank L. and John E. Hunter. 1983. "Individual Differences in Productivity: An Empirical Test of Estimates Derived from Studies of Selection Procedure Utility," *Journal of Applied Psychology* Vol 68, pp. 407-414.
- Zoega, Gylfi and Alison L. Booth. 2005. "Worker Heterogeneity, Intra-Firm Externalities, and Wage Compression." Birkbek Working Papers in Economics and Finance, Number 0515.

Table 1

Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Ln Prod. Growth over initial 2 years	0.80	1.14	-1.10	6.91
Ln Prod. Growth over initial 3 months	0.51	0.94	-1.37	6.69
Ln Prod. Growth from third month until end of second year	0.29	0.43	-1.25	6.22
Ln Wage Growth	0.18	0.20	-1.99	1.95
Age	27.00	9.08	16.00	64.00
Vocational schooling	0.28	0.45	0.00	1.00
Temporary or seasonal job	0.15	0.35	0.00	1.00
Part-time job	0.21	0.41	0.00	1.00
Union	0.11	0.28	0.00	1.00
Ln establishment size	2.91	1.51	0.00	8.60
Female	0.45	0.50	0.00	1.00
Tenure (weeks)	67.38	71.55	1.10	1152.67
High School Indicator	0.76	0.42	0	1
College indicator	0.12	0.32	0	1
Ln (formal tr. hours+ 1)	0.48	1.30	0.00	6.46
Ln (informal tr. hours + 1)	3.43	1.52	0.00	7.60
Ln (hours watch others, typical, + 1)	2.75	1.73	0.00	6.88
Ln (weeks until fully trained, typical, + 1)	2.21	1.23	0.00	6.03
Ln (weeks rel. experience + 1)	2.86	2.44	0.00	7.64
Ln (formal tr. , typical + 1)	0.50	1.34	0	6.69
Ln (informal tr. , typical, + 1)	3.56	1.48	0	7.60
Obs.*	1, 543			

*There are 1,471 observations with information on relevant experience and 1,538 observations with information on informal training.

Table 2

OLS Wage Equations

	Starting Productivity is average productivity during the first two weeks	Starting Productivity is average productivity during the first three months
Typical worker's log productivity growth over two years	0.027 (0.004)	0.073 (0.011)
Difference between log starting productivity of typical worker and log starting prod. of most recent hire	0.015 (0.008)	0.023 (0.013)
Union	-0.011 (0.018)	-0.010 (0.018)
Log of establishment size	0.0005 (0.0036)	0.0013 (0.0036)
Weeks tenure	0.0009 (0.0001)	0.0009 (0.0001)
High School difference*	0.007 (0.016)	0.007 (0.016)
College difference	0.036 (0.021)	0.035 (0.021)
Vocational training difference	-0.001 (0.011)	0.006 (0.011)
Female difference	0.003 (0.010)	0.003 (0.010)
Age difference	0.009 (0.003)	0.009 (0.003)
Age squared difference/100	-0.01 (0.00)	-0.01 (0.00)
Temporary difference	0.002 (0.014)	0.005 (0.014)
Part-time difference	0.005 (0.014)	0.006 (0.013)
Constant	0.010 (0.014)	0.095 (0.014)
Number of Observations	1,543	1,543
\bar{R}^2	.13	.13

*High School difference is the high school indicator for the typical worker minus the high school indicator for the most recent hire. The other dif variables in the table are defined analogously.

Table 3

Selected Coefficients, Productivity Growth Equation

Ln (weeks rel. experience + 1)	0.032 (0.007)
Ln (formal tr. hours + 1)	-0.16 (0.10)
Ln (informal tr. hours + 1)	-0.23 (0.04)
Ln (weeks until fully trained, typical, + 1)	-0.15 (0.03)
Ln (hours watch others, typical, 1 + 1)	-0.13 (0.03)
Ln (formal tr. hours, typical, + 1)	0.23 (0.10)
Ln (informal tr. hours, typical, + 1)	0.12 (0.04)
Ln (formal tr. hours + 1) x (Tenure spline 1)	0.069 (0.015)
Ln (informal tr. hours + 1) x (Tenure spline 1)	0.045 (0.014)
Ln (weeks until fully trained, typical, + 1) x (Tenure spline 1)	0.049 (0.012)
Ln (watch others, typical, + 1) x (Tenure spline 1)	0.049 (0.012)
Ln (formal tr. hours, typical, + 1) x (Tenure spline 1)	0.018 (0.035)
Ln (informal tr. hours, typical, + 1) x (Tenure spline 1)	-0.36 (0.02)
Ln (formal tr. hours + 1) x (Tenure spline 2)	0.009 (0.028)
Ln (informal tr. hours + 1) x (Tenure spline 2)	0.004 (0.015)
Ln (weeks until fully trained, typical, + 1) x (Tenure spline 2)	0.008 (0.018)
Ln (watch others, typical, + 1) x (Tenure spline 2)	0.008 (0.013)
Ln (formal tr. hours, typical, + 1) x (Tenure spline 2)	-0.024 (0.026)
Ln (informal tr. hours, typical, + 1) x (Tenure spline 2)	0.000 (0.014)
Number of Observations	4,446
R ²	0.21

Tenure spline 1 = $\min(\ln(t+1), \ln(12+1))$

Tenure spline 2 = $\max(\ln(t+1) - \ln(12+1), 0)$

Table 4

IV Estimation of the Sharing Coefficients*

Contract Length	Difference between log starting productivity of typical worker and log starting prod. of most recent hire (b_1)	Difference between prod. growth of typical worker and prod. growth of most recent hire (c_1)	Productivity growth of typical worker over the two year period (c_2)
12 weeks	0.32 (0.07)	0.19 (0.29)	0.19 (0.03)
18 weeks	0.32 (0.08)	0.12 (0.41)	0.33 (0.08)
26 weeks	0.32 (0.08)	0.02 (0.43)	0.46 (0.16)
38 weeks	0.32 (0.09)	-0.13 (0.49)	0.58 (0.24)
52 weeks	0.31 (0.09)	-0.43 (0.58)	0.61 (0.26)
Variable	0.32 (0.10)	0.17 (0.46)	0.26 (0.05)
n = 1518			

* Standard errors generated by bootstrapping.

Table 5

Probit Equation: Probability of a Wage Change

Typical worker's log productivity growth over the first two years	0.57 (0.41)
Typical worker's log productivity growth x weeks tenure	0.012 (0.007)
Typical worker's log productivity growth x (weeks tenure) ² /100	-0.0013 (0.0008)
Weeks tenure	0.064 (0.025)
(Weeks tenure) ² /100	-0.068 (0.037)
Tenure >= 1 year	1.18 (0.41)
Weeks tenure x (Tenure >= 1 year)	-0.058 (0.024)
(Weeks tenure) ² x (Tenure >= 1 year)/100	0.067 (0.037)
Constant	-1.21 (0.49)
Number of Observations	1686