Wage Dispersion, Compensation Policy, and the Role of Firms[1]

Bryce Stephens[2]

University of Maryland

stephens@econ.umd.edu

JOB MARKET PAPER

November 20, 2005

[1]This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01 AG018854-02, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau supports external researchers' use of these data through the Research Data Centers (see www.ces.census.gov). For other questions regarding the data, please contact Jeremy S. Wu, Director, U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA. (Jeremy.S.Wu@census.gov http://lehd.dsd.census.gov).

[2]I would like to thank John Haltiwanger, John Abowd, and John Shea for their helpful comments, advice, and support. I would also like to thank participants in the macroeconomics brown bag seminars at the University of Maryland as well as Kevin McKinney, Erika McEntarfer, Martha Stinson, Lars Vilhuber, Simon Woodcock, and other staff of the LEHD program.

Abstract

Empirical work in economics stresses the importance of unobserved firm- and person-level characteristics in the determination of wages, finding that these unobserved components account for the overwhelming majority of variation in wages. However, little is known about the mechanisms sustaining these wage differentials. This paper attempts to demystify the firm-side of the puzzle by developing a statistical model that enriches the role that firms play in wage determination, allowing firms to influence both average wages as well as the returns to observable worker characteristics.

I exploit the hierarchical nature of a unique employer-employee linked dataset for the United States, estimating a multilevel statistical model of earnings that accounts for firm-specific deviations in average wages as well as the returns to components of human capital – race, gender, education, and experience – while also controlling for person-level heterogeneity in earnings. These idiosyncratic prices reflect one aspect of firm compensation policy; another, and more novel aspect, is the unstructured characterization of the covariance of these prices across firms.

I estimate the model's variance parameters using Restricted (or Residual) Maximum Likelihood techniques. Results suggest that there is significant variation in the returns to worker characteristics across firms. First, estimates of the parameters of the covariance matrix of firm-specific returns are statistically significant. Firms that tend to pay higher average wages also tend to pay higher than average returns to worker characteristics; firms that tend to reward highly the human capital of men also highly reward the human capital of women. For instance, the correlation between the firm-specific returns to education for men and women is 0.57. Second, the firm-specific returns account for roughly 9% of the variation in wages – approximately 50% of the variation in wages explained by firm-specific intercepts alone. The inclusion of firm-specific returns ties variation in wages, otherwise attributable to firm-specific intercepts, to observable components of human capital.

Keywords: compensation policy, employer-employee longitudinal data, human capital

JEL Codes: C23, J24, J30, J31, M50

1 Introduction

Empirical work in economics stresses the importance of unobserved firm- and person-level characteristics in the determination of wages, finding that these unobserved components account for the overwhelming majority of variation in wages. However, little is known about the mechanisms sustaining these wage differentials. This paper attempts to demystify the firm-side of the puzzle by developing a statistical model that enriches the role that firms play in wage determination, allowing firms to influence both average wages as well as the returns to observed worker characteristics.

A burgeoning literature exists focused on measuring employer wage differentials. In nearly all of these studies, firm wage differentials are measured as a firm-specific intercept: Groshen (1991) establishes the importance of establishment wage differentials relative the inter-industry wage differentials; Groshen & Levine (1998) measure the magnitude and persistence of firm- and occupation-wage differentials in an attempt to assess the importance of internal labor markets; Bronars & Famulari (1997) provide evidence of employer wage differentials in their investigation of the determinants of wage levels and wage growth; while Abowd, Kramarz & Margolis (1999) use employer-employee connected data to implement a technique for estimating a wage model with both firm and person effects. These differentials are often decomposed into the portion explained by observed firm characteristics, for example, industry of operation, and an unobserved component – the portion of wages explained by firm identity that is not explained by firm characteristics available to the econometrician. A similar decomposition is applied to the individual component of wages. Person-level intercepts – the time invariant portion of wages attributed to individuals – are decomposed into a component explained by time invariant, though observed, characteristics such as race or sex and a component that is unobserved. The person-specific unobserved portion of wages is often thought of as reflecting the value of an individual's innate ability or talent that is portable across firms.

A complementary literature explores the relationship between firm characteristics and wage outcomes. For example, empirical work suggests that there exists a significant relationship between wages and firm size. Brown & Medoff (1989) provides an analysis of the role of firm size in the determination of wages. Large firms may pay higher wages because large firms hire higher quality workers, offer inferior working conditions, are more threatened by unionization, or face higher monitoring costs (e.g., efficiency wages). Davis & Haltiwanger (1995), using data from the Current Population Survey and Census of Manufacturers, investigate the relationship between firm size and both wage levels and wage dispersion. They find that average wages are higher at larger establishments and that wage dispersion is inversely related to firm size class. In addition, their work suggests that the returns to observed characteristics vary across firm size class.

Finally, economists are also looking within the firm to explain wage setting from the perspective of compensation and strategic management policies. As Lazear (2000) mentions in his review of the personnel economics literature, firms may influence wage determination and wage dynamics through firm level policies by designing deferred compensation policies to motivate workers, tying remuneration to observed productivity, or designing tournament schemes. Firms may also pay efficiency wages to dissuade shirking and may provide insurance against shocks to the value of labor services (Prendergast 1999). Recent work in Ichniowski & Shaw (2003) and Cappelli & Neumark (2001) evaluates the relationship between firm-level human resource practices and firm productivity and wage outcomes. Finally, the work of Baker, Gibbs & Holmstrom (1995) illustrates the value in evaluating the wage policies of an individual firm.

The empirical work focused on decomposing wages, emphasizing the relative importance of the returns to observed characteristics and unobserved heterogeneity, has in some sense preceded the evolution of economic theory that explains why unobserved heterogeneity exists. Firms pay vastly different wages, though, observed firm characteristics fall short in explaining the overwhelming majority of the variation in average wages across firms. As economists begin to look inside the firm for evidence supporting this measured heterogeneity, they are finding that firms pursue idiosyncratic compensation policies and human resource management strategies. The work in this paper pushes this investigation one step further, suggesting that digging deeper into the black box of firm compensation policy requires evaluating not only traditional firm- and person-specific determinants of wages, but also the way in which the characteristics of workers are valued by firms. Heterogeneity in firms' policies, which may explain why firms pay different average wages, may also suggest that certain worker attributes are more valuable to particular firms.

Using employer-employee linked data for the United States, I specify and estimate a multilevel statistical model of earnings determination that measures variation in firm wage policies. The specification controls for person-level, unobserved heterogeneity and allows for firm-specific deviations in average firm wages as well as in the returns to components of human capital: race, gender, education, and experience. The estimation procedure also provides estimates of the elements of the variance matrix of pay policy parameters. Results suggest that there is statistically significant variation in the returns to worker characteristics across firms. For instance, the estimated variance of the returns to being non-white, for both men and women, is roughly the same size as the variance of average wages across firms. Allowing for an unstructured covariance matrix of firm-specific returns reveals that firms that tend to pay higher average wages also tend to pay higher than average returns to certain worker characteristics. Finally, roughly 9% of wage variation is accounted for by the firm-specific returns to human capital – approximately 50% of the variation in wages explained by firm-specific intercepts alone. Though it appears that the majority of this explained variation would have otherwise – in the absence of firm-specific returns – been accounted for by person- and firm-specific intercepts, firm-specific valuation of human capital explains a greater proportion of wage variation than the returns to the observed firm characteristics included in the estimation: the level of firm employment and industry division of operation.

Hierarchical modeling techniques are used in Cardoso (2000) to analyze the relationship between firm level characteristics and worker wages in a cross-section of employer-employee connected data for Portugal. The model presented here extends the approach in Cardoso (2000) by integrating it into a larger literature that accounts for unobserved heterogeneity across workers and firms, by permitting a fully unstructured variance matrix of pay policy parameters, and by accounting for the total variation in wages attributed to the firm-specific returns to human capital. My model may also be viewed as an extension of Abowd et al. (1999) in that it allows unobserved heterogeneity across firms to be captured by firm-specific intercepts as well as firm-specific returns. By exploiting the longitudinal and connected nature of the employer-employee dataset, I am able to control for both unobserved firm- and person-effects. I use Restricted Maximum Likelihood (REML) techniques to estimate the model variance and covariance parameters and derive random coefficient estimates using the Henderson (or Mixed) Model equations. Unlike traditional approaches to estimating random effects models, for instance, Generalized Least Squares, the REML approach does not assume orthogonality between the random effects – the firm-specific slopes and firm- and person-specific intercepts – and observed covariates. My approach also allows for a fully unstructured variance matrix of the firm-specific intercepts and returns. Finally, I provide a decomposition of the total variation in wages attributed to both the observed and unobserved components in the model, thereby highlighting the relative importance of the firm-specific components of compensation.

The paper is organized as follows. Section 2 discusses the multilevel model in the context of firm compensation policy; Section 3 develops the statistical model and discusses the estimation technique; Section 4 provides a description of the data; Section 5 provides results; and Section 6 concludes.

2 Firm Compensation Policy

To evaluate the influence of firms on wages, I specify a multilevel (mixed) model of wage determination. The multilevel characteristic of the model captures the inherently hierarchical nature of the data: individuals hold jobs (level 1) that are nested within firms (level 2). The model is mixed in that it contains both random and fixed coefficients. Firms shape individual wage outcomes in two ways: through firm-specific average wage effects and firm-specific price effects. Average wages are influenced by firm-specific intercepts as well as by observed characteristics, namely, firm employment (size) and industry division of operation. Firms also influence wage setting through the returns to the characteristics of human capital: sex, race, education, and experience.[1] These returns have an unobserved firm-specific component, measured as deviations around the sample average return for a given characteristic of human capital, as well as an observed component − the interaction between observed characteristics of workers and firms. For example, consider the influence of firms on wage setting when only firm size and experience are observed. Firms will influence wages through a firm-specific average wage effect that has two components: an unobserved, though identifiable, firm-specific random intercept and the observed effect of firm size. Firms will also influence wages through the returns to experience, which also has two components: an unobserved, though identifiable, firm-specific random return to experience and the observed interaction between experience and firm size.

The statistical model retains characteristics of more standard models of wage determination. Workers earn average (or market) returns to the components of human capital which are measured as fixed coefficients on the returns to human capital. Thus, the firm-specific random returns may be viewed as deviations around the sample average return to the components of human capital. A person-specific, random intercept is also introduced to control for unobserved heterogeneity and is usually interpreted as the return to a worker's innate ability or skill. Finally, a number of dummy variables are included to control for time effects and issues related to the construction of the data.[2]

The econometric specification, while driven partly by the structure of the data, is loosely supported by theoretical work in two recent papers. In Abowd, Kramarz & Roux (2005), a simple model of production, wage determination, and mobility is specified. There are two types of firms: complex and simple. In simple firms, wages are set equal to a worker's productivity which is determined by her innate ability and level of experience. Productivity in complex firms, however, is firm-specific and is a function of an individual's innate ability and firm-specific seniority. Wages are determined by a simple sharing rule and, for workers at complex firms, wages are thus a function of tenure at the firm. The underlying model supports an empirical specification in which a firm-specific wage policy captures both the sharing rule and a production technology that is a function of tenure at the firm. The model, while particularly relevant to the literature on mobility and the returns to tenure, suggests a broader treatment of firm specificity in production and the formation of compensation policy. Lazear (2003) provides another simple and more general model of output and wage determination. A worker's output is determined by a set of general skills which are valued by all firms in the economy. Firms differ, however, in the relative weighting of these skills in production. Workers and firms split the value of output and a worker's earnings are determined by his share of output − a function of his optimal investment in general skills and the firm's specific weighting of his skills. The model is primarily developed to provide an explanation for the empirically observed returns to firm-specific tenure,

[1] In the implementation, the model is fully interacted by sex; firm-specific returns are estimated for both men and women.

[2] These variables are discussed in further detail in Section 4. A list of variables included in the estimation is presented in Table 1.

though is relevant to the statistical model presented in this paper which allows for firm-specific returns to the components of human capital that are used by all firms. In light of the model developed in Lazear (2003), the firm-specific returns to human capital in this paper may be viewed as capturing the firm-specific weighting of these characteristics in production.

Efficiency wage theory, which generally refers to extra-marginal wage payments at the firm-level, may also support a model of firm-specific returns. Firms that pay high average wages to all of their workers may do so to provide a disincentive to shirking. It may be, however, that the level of certain worker attributes – experience and education – make monitoring by firms more difficult. More experienced workers may be more likely to shirk because, over their careers, they have learned to shirk successfully. It may also be that workers with more experience and education are more likely to have more complex jobs where output is hard to observe and, thus, hard to monitor. In either case, one might expect to see higher than average firm-level returns to these characteristics where monitoring is difficult. If the existence of high average firm wages suggests that firms are paying efficiency wages, then it seems reasonable – if the monitoring of highly educated and experienced workers is more difficult – to expect a positive covariance between firm-specific average wages (intercepts) and firm-specific returns to education and experience.

Regardless of theoretical motivation for the multilevel specification, the structure of the model assumes that worker-firm attachments, i.e., the clustering of workers within firms, is meaningful. Econometrically, the treatment of firm-specific deviations as random effects provides an efficiency gain over other modeling approaches that do not take into account the clustering of workers within firms in cases when this clustering is present in the data.[3] A richer error specification will provide correct standard errors for the fixed effects in the model. Random firm-specific effects identified in the data, however, may exist for a variety of reasons. For instance, consider the specific return to education. If the measure of education is not quality adjusted, which is the case in the data used in this analysis, a higher firm-specific return may be due to the clustering of workers with high levels of education from high quality universities who earn high wages relative to workers at the same firm who have lower levels of education from low quality schools. Dispersion in the returns to education, thus, would result from the inability to observe the quality of education. Econometrically, accounting for this clustering is meaningful in the aforementioned sense. However, if the force driving dispersion in the returns to education is due to unobserved quality issues, the economic content of these slopes is questionable. It may be, however, that dispersion in returns across firms are meaningful in economic terms. In the previous example, if firms that hire the highest quality university graduates also hire the highest quality high school graduates, then the existence of higher than average firm-level returns to education may reflect a hiring policy: always hire the highest quality of graduates regardless of level

[3] An example is Ordinary Least Squares (OLS), where the model error is assumed spherical.

of education. As will be made clear in the development of the model, these firm-specific deviations may be thought of as the firm-specific unobserved error components in a pricing model of the characteristics of human capital. From this perspective, the existence of firm-specific returns is due to the inability of the observed firm characteristics in the model to explain variation in the returns to the characteristics of human capital. In either case, the approach permits an accounting for the contribution of observed and unobserved person and firm characteristics to variation in wages.

In addition to allowing for firm-specific intercepts and returns, the multilevel model is flexible in accounting for both the variance and covariance of these firm-specific returns. For each firm-specific component, a sample-wide variance is specified.[4] The structure of the covariance of these effects, however, may take a variety of forms. In this analysis, I evaluate both a diagonal and fully unstructured variance matrix of firm effects. Exploring the unstructured form of the of this matrix addresses potentially important questions regarding firm wage policy. For instance, do firms that tend to pay high wages on average also tend to reward more highly the returns to education? Is the dispersion in the returns to education across firms similar for men and women?

The model also includes a person-specific random intercept. Though jobs are nested within firms, the person-specific random intercept exists outside the hierarchy that is developed in the following section. It is included to capture some of the unobserved heterogeneity in wages that is due to what is traditionally considered an individual's innate ability or skill and also to serve as a basis for comparison to work that models heterogeneity in wages with both firm- and person-specific intercepts. Future work will attempt to integrate more fully the person-specific intercepts into the variance structure of the multilevel model. For instance, permitting the covariance between the person-specific intercept and the firm-specific returns to education would provide a direct empirical test of whether individuals with high levels of innate ability (higher than average person-intercepts) are more likely to earn higher returns to experience.

A thorough discussion of multilevel models is provided in Goldstein (1995). Raudenbush & Bryk (1986) also provide a discussion and application to estimating the returns to student achievement. Cardoso (2000) and Cardoso (1999) develop a multilevel wage model using employer-employee connected data for Portugal that is estimated using Iterative Generalized Least Squares (IGLS) techniques that are popular in the multilevel model literature. The multilevel model, which includes both fixed and random coefficients, is also essentially a mixed model. Estimation of and prediction in mixed models is discussed at length in McCulloch & Searle (2001) and Searle, Casella & McCulloch (1992). The approach in this paper draws heavily on these techniques.

[4] Each effect is assumed to have a zero mean. In the analysis, the data are de-meaned so that disperison of the firm interecpts is appropriately centered. For the firm-specfic slopes, the fixed coefficient estimates of the returns to the components of human capital provide estimates of the mean of each distribution.

A general multilevel model is outlined in the following section and is shown to have a mixed model representation. Also discussed is the three-part estimation approach: (1) the natural log of the real wage, the dependent variable, is transformed, removing variation attributed to the variables for which coefficients are fixed; (2) Restricted Maximum Likelihood (REML) techniques are used to estimate the variance components of the model; and (3) predictors of the model's fixed and random effects are derived using the Henderson (or Mixed Model) equations. Attention is given to issues surrounding the identification of the model's random effects and issues involved in the identification of these effects if they were assumed fixed.

3 Statistical Model

3.1 Multilevel Model

Let $i = 1, ..., N$ index workers, $j = 1, ..., J$ index firms, and $t = 1, ..., T$ index time. The natural log of the real wage, w_{ijt}, for worker i employed by firm j in time t is modeled as:

$$w_{ijt} = x_{ijt}^{(1)\prime}\beta^{(1)} + x_{it}^{(2)\prime}\gamma_{jt} + \alpha_i + \varepsilon_{ijt} \tag{1}$$

$$x_{ijt}^{(1)} = \begin{bmatrix} x_{1ijt}^{(1)} \\ \vdots \\ x_{mijt}^{(1)} \end{bmatrix}, x_{ijt}^{(2)} = \begin{bmatrix} x_{1ijt}^{(2)} \\ \vdots \\ x_{kijt}^{(2)} \end{bmatrix}$$

where $x_{ijt}^{(1)}$ is an $(m \times 1)$ vector of person and firm varying covariates for which parameters $\beta^{(1)}$an $(m \times 1)$ vector are fixed;[5] $x_{it}^{(2)}$ is a $(k \times 1)$ vector of person-level components of human capital (including an intercept) for which parameters γ_{jt} a $(k \times 1)$ vector are firm-specific;[6] $\alpha_i \sim N(0, \sigma_\alpha^2)$ is a person-level random intercept; and $\varepsilon_{ijt} \sim N(0, \sigma_\varepsilon^2)$ is a residual error term.

The second level of the model involves specifying a relationship between γ_{jt} – the vector of firm-specific returns to human capital – and firm characteristics:

$$\gamma_{jt} = g_{jt}'\beta^{(2)} + \Psi_j \tag{2}$$

where:

[5] The elements of $x_{ijt}^{(1)}$ are: a race missing dummy, negative experience dummy, and time effects interacted with sex. The superscript (1) signifies level 1 covariates for which parameters do not vary over the second level.

[6] The elements of $x_{it}^{(2)}$ include: a constant, experience, $\frac{experience^2}{100}$, education, and a nonwhite dummy, all of which are interacted with sex. The superscript (2) is to signify level 1 covariates for which parameters do vary over the second level.

$$\underset{(kl\ x\ k)}{g_{jt}} = I_k \otimes f_{jt} \tag{3}$$

f_{jt} is an (l ×1) vector of time-varying firm covariates – firm-level employment and industry of operation:

$$f_{jt} = \begin{bmatrix} f_{1jt} \\ \vdots \\ f_{ljt} \end{bmatrix} \text{ where } f_{1j} = 1, \Psi_j = \begin{bmatrix} \psi_{1j} \\ \vdots \\ \psi_{kj} \end{bmatrix} \tag{4}$$

$\beta^{(2)}$ is a (kl ×1) vector of fixed parameters describing the relationship between these firm characteristics and the returns to person-level covariates and average wages. Ψ_j is a (k ×1) vector of firm-specific random errors with each element of Ψ_j corresponding to each of the k components of human capital. Thus, firms influence the returns to human capital in two ways. The term $g'_{jt}\beta^{(2)}$ reflects the influence of time varying, observed firm characteristics. For example, the firm-specific returns to education, in the context of this model, are influenced by both the size and industrial classification of the employing firm. This component will capture whether firms that tend to be large in terms of employment also tend to pay higher returns to experience. The vector Ψ_j, the firm-specific error component, captures the influence of the unobserved, firm-specific component of compensation policies. All variation across firms in the return to a component of human capital that is not captured by $g'_{jt}\beta^{(2)}$ is accounted for by the elements of Ψ_j. This error representation, for example, would capture the existence of firm-specific human resource or management policies that would otherwise not be captured by firm observed characteristics. Finally, it is important to notice that while the firm-specific parameters γ_{jt} vary over time, this variation is due to time variation in the value of f_{jt} and not Ψ_j. Thus, stochastic changes in the firm-specific error components are smoothed over time.[7]

A characteristic of the model is that the vector of firm-specific errors Ψ_j is permitted to have a fully unstructured variance matrix. It is assumed that $\Psi_j \sim N(0, \Gamma)$ where:

$$\Gamma = \begin{bmatrix} \sigma^2_{\psi_1} & \cdots & \sigma_{\psi_1\psi_k} \\ \vdots & \ddots & \vdots \\ \sigma_{\psi_1\psi_k} & \cdots & \sigma^2_{\psi_k} \end{bmatrix} \tag{5}$$

The characterization of Γ in equation (5) is general, capturing the variances of the firm-specific intercepts and returns *as well as* the covariances between these terms. For instance, $\sigma^2_{\psi_1}$ is the variance of the sample-

[7] This will turn out to be important as the components of Ψ_j will later be interpreted as firm-specific returns to human capital. Increases in the returns to these components – for example, due to increases in the returns to skill – will be smoothed over the 1990-1998 sample used in the analysis.

wide firm-specific intercepts, measuring the dispersion in wages due to dispersion in firm-specific average wages. The extent to which firms in sample tend to pay high average wages and high returns to education, experience, and race for men and women will be captured by covariance terms in the first column (or first row) of Γ.

Combining equations (1) and (2) yields:

$$w_{ijt} = x_{ijt}^{(1)\prime}\beta^{(1)} + x_{it}^{(2)\prime}g_{jt}^{\prime}\beta^{(2)} + x_{it}^{(2)\prime}\Psi_j + \alpha_i + \varepsilon_{ijt} \tag{6}$$

The influence of the firms (second level of the model) is readily seen in equation (6). The first term on the right hand side of equation (6) captures the influence of the person-level covariates for which parameter variation is not firm-specific – the set of time and control variables. The second term captures the influence of observed, firm-level characteristics (influencing variation in γ_{jt}) on the components of human capital. Recalling the structure of g_{jt} in equation (2), the multiplication of $x_{it}^{(2)\prime}g_{jt}^{\prime}$ yields a vector containing the components of human capital, firm observed characteristics, and the interactions between the components of human capital and observed firm characteristics. The returns to the components of human capital in $x_{it}^{(2)\prime}g_{jt}^{\prime}\beta^{(2)}$ are viewed as the market or sample-wide average returns to these characteristics; the returns to observed firm characteristics and the interactions capture the influence of the observed firm characteristics in equation (2). The firm-specific random errors Ψ_j in equation (2), in the context of equation (6), become a set of firm-specific random coefficients for the components of human capital.

In order to explain the technique used to estimate the model, it is helpful to express the model in matrix notation. Doing this takes a bit of finessing. Express $x_{ijt}^{(1)\prime}\beta^{(1)} + x_{it}^{(2)\prime}g_{jt}^{\prime}\beta^{(2)}$ in (6) as:

$$x_{ijt}^{(1)\prime}\beta^{(1)} + x_{it}^{(2)\prime}g_{jt}^{\prime}\beta^{(2)} = b_1^{(1)}x_{1ijt}^{(1)} + \cdots + b_m^{(1)}x_{mijt}^{(1)} + b_1^{(2)}x_{1it}^{(2)}g_{1jt} + \cdots + b_{kl}^{(2)}x_{kit}^{(2)}g_{ljt} = x_{ijt}^{\prime}\beta$$

where x_{ijt} is a $((kl+m)\ \times 1)$ vector of covariates – those contained in $x_{ijt}^{(1)}$ which are specified as having fixed coefficient estimates in equation (1) and the interactions of the $x_{ijt}^{(2)}$ in equation (1) with the firm-level characteristics f_{jt} in equation (4). β is a $((kl+m)\ \times 1)$ vector of fixed coefficients. Equation (6) becomes:

$$w_{ijt} = x_{ijt}^{\prime}\beta + x_{it}^{(2)\prime}\Psi_j + \alpha_i + \varepsilon_{ijt} \tag{7}$$

Let N index the number of observations, I index the number of persons, and J index firms. Grouping observations into firms and stacking yields:

$$w = X\beta + X^{(2)}\Psi + D\alpha + \varepsilon \tag{8}$$

where w is an $(N \times 1)$ vector of earnings; X is an $(N \times (kl+m))$ matrix of covariates; and β is a $((kl+m) \times 1)$ vector of fixed coefficients. $X^{(2)}$ is an $(N \times Jk)$ stacked matrix of person-level characteristics (the $x_{it}^{(2)\prime}$ grouped by firm):

$$X^{(2)} = \begin{bmatrix} X_1^{(2)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & X_J^{(2)} \end{bmatrix} \tag{9}$$

so that $X_j^{(2)}$ is an $(n_j \times k)$ matrix of observations of workers at firm J. Ψ is a $(Jk \times 1)$ vector of firm specific random returns:

$$\Psi = \begin{bmatrix} \Psi_1 \\ \vdots \\ \Psi_J \end{bmatrix} \tag{10}$$

and is ordered so that the firm specific returns in Ψ_j (recall that Ψ_j is a $(k \times 1)$ vector) correspond appropriately to the block of observations in $X^{(2)}$ for that firm. Formally, $X^{(2)}$ is the design of firm-specific random effects and includes firm-specific random intercepts and returns to worker level characteristics. Finally, D is an $(N \times I)$ design matrix for the random person effects contained in the $(I \times 1)$ vector α and ε is an $(N \times 1)$ vector of residuals.

The model's error is represented in matrix notation as follows:

$$\begin{bmatrix} \Psi \\ \alpha \\ \varepsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_J \otimes \Gamma & 0 & 0 \\ 0 & \sigma_\alpha^2 I_I & 0 \\ 0 & 0 & \sigma_\varepsilon^2 I_N \end{bmatrix} \right) \tag{11}$$

While the variance of the firm-specific returns and intercepts is fully unstructured – the term $I_J \otimes \Gamma$ – and a structure is placed on the person-specific intercepts $\sigma_\alpha^2 I_I$, the person-specific intercepts α are not permitted to co-vary with the elements of Ψ. This is a restrictive assumption. While the thrust of this paper is to explore a specific characterization of firm compensation policy – namely, one in which firms are permitted to pay specific returns and to adjust these parameters such that high returns to some components may be associated with high or low returns to others – it is reasonable to believe that individuals with higher than average levels of innate ability, measured by relatively high α's, may also earn higher returns to certain characteristics or may be more likely to match to firms that pay higher than average wages. A context in which the current error structure in (11) seems reasonable is one where firms are unable to implement compensation policies conditional on innate (and initially unobserved) worker ability and where learning about innate ability, through employment at the firm, does not induce adjustments in firm compensation policy.

3.2 Model Estimation and Prediction

Methods for estimating the model specified by equations (8) and (11), in the context of the multilevel model literature, are discussed in Goldstein (1995) and include Iterative Generalized Least Squares (IGLS), Bayes or Empirical Bayes, and Generalized Estimating Equations (GEE) methods. Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML) are briefly mentioned as well. Estimation of the model has also received considerable treatment in the statistics literature. Equations (8) and (11) represent a mixed model formulation as the specification contains both random and fixed coefficients.

The specifications in this paper are estimated using REML, which provides estimates of the variance components in (11): $\widehat{\sigma}_{\alpha}^{2}$, $\widehat{\sigma}_{\varepsilon}^{2}$, and the elements of $\widehat{\Gamma}$. Loosely speaking, "a basic idea of restricted maximum likelihood (REML) estimation is that of estimating variance components based on residuals calculated after fitting by ordinary least squares just the fixed effects part of the model" rather than basing these estimates on the dependent variable (Searle et al. (1992) p. 250). See Appendix A for a more detailed discussion of the REML technique.

Using the variance estimates emerging from REML, predictors of the random coefficients – $\widehat{\Psi}$ and $\widehat{\alpha}$ – and estimates of the fixed coefficients $\widehat{\beta}$ in equation (8) are derived from the Mixed Model (or Henderson) Equations. See Appendix B for a full derivation of the predictors and estimates as well as a discussion of their identification.

4 Data Description

I estimate the multilevel model using data that are house at the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Program. A thorough discussion of the data maintained by LEHD is provided in Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer & Woodcock (2005); a brief description is provided here. The LEHD Program maintains a variety of survey and administrative data from a number of federal and state agencies. State unemployment insurance (UI) wage records and Quarterly Census of Employment and Wages (QCEW) establishment data are available for the states in partnership with the LEHD Program.[8] State-level unemployment insurance (UI) data contain quarterly earnings for employees covered by state unemployment insurance systems (over 96% of private sector employment is covered by the

[8] Currently, 40 states have signed Memoranda of Understanding to engage in data sharing with the U.S. Census Bureau. For 30 of these states, core infrastructural and public use data are available. More information on state partnerships with the U.S. Census Bureau's LEHD Program is available at: http://lehd.dsd.census.gov.

UI universe) as well as both a person identifier – a person-specific Protected Identification Key (PIK) – and a firm identifier – a State Employer Identification Number (SEIN). The structure of the UI data conforms to the data requirements of the multilevel model specifications: there are multiple workers in the UI data for whom wages are reported within the same firm (workers are clustered within firms). Thus, a firm, as defined in this analysis, is a collection of workers who share a common SEIN.[9] And, individual wage records are then linked across quarters on the basis of the PIK to create individual work histories.

Nearly all of the firm identifiers in the UI data are found in the universe of state QCEW data which contain richer information on firms. However, while groupings of workers by a common SEIN define the smallest firm in the UI data, a finer definition of the firm is available in the QCEW data.[10] For SEINs in the QCEW data that operate more than one establishment, firm characteristics such as employment, total payroll, location, and industry of operation are reported at the level of the establishment. While the finer, establishment-level detail is not used in this analysis, the QCEW do provide the measure of industry of operation. For firms in the QCEW reporting under multiple establishments (a multi-unit firm), the SEIN-level industry is the employment weighted modal industry of the underlying establishments. For firms reporting under a single establishment, the reported industry is used.

Finally, worker demographics – sex, date of birth, and race – are acquired from the Census NUMIDENT file (a version of Social Security Administration person-level micro-data) and are matched to the UI wage record data on the basis of the person identifier. Neither the UI nor the Census NUMIDENT data provide a direct measure of education for individuals in the sample. Education information is available, however, in the Decennial Census of Population; in this analysis, the education variable for workers in the UI data is based on a statistical match to the 1990 Decennial Census of Population.[11] The race variable is collapsed into an indicator for non-white, rather than used directly in the model estimation.[12]

All worker and firm measures are transformed into annual values. From the QCEW data, only industry of operation is retained in this analysis. For each firm (SEIN), the annual (calendar year) employment weighted, quarterly modal industry is used.[13] The measure of firm size is simply the summation of workers with positive earnings at the firm within the calendar year. Date of birth from the NUMIDENT is used to create an age measure which, in turn, is used to create a measure of potential experience. For the first

[9] The SEIN is by definition a state-level firm identifier; thus, firms are also state-specific.

[10] This is true for all states except Minnesota. The Minnesota UI data provide two firm identifiers – an SEIN and reporting unit number – which are consistent with the level of reporting in the QCEW.

[11] Work is currently being done at LEHD to match workers in the UI wage data to the 2000 Decennial Census of Population on the basis of PIK, thus, providing reported education measures for roughly 1 in 6 workers in the UI universe. For remaining workers in the UI data, education levels will be created using multiple imputation techniques.

[12] The collapsing of the race variable is common in nearly all of the wage research using LEHD data. Preliminary research into the quality of the race variable on the NUMIDENT raised concerns regarding it's usefulness in analysis. In the future, I plan to explore the reliability of using the race variable in the context of the model developed in this paper.

[13] Firm data in the QCEW are reported on a quarterly basis. While it is rare for firms to change industry classification, the algorithm ensures that the most important industry – in terms of employment – is used in the analysis.

year that an individual appears in the UI data, potential experience is calculated as the person's age at the beginning of the quarter minus her years of education minus 6. This potential experience measure is then augmented by observed years of experience: years during which the individual has positive earnings in the UI data.

The wage measure used in the analysis is based on an annualization of the quarterly payroll values reported in the UI data. For workers in the UI data who hold multiple jobs concurrently, the dominant job with the highest level of annual earnings is retained. Within each dominant job, quarterly earnings within the calendar year are used to create a measure of annualized earnings (wages). The sequence of earnings within a job are used to categorize workers into three groups which are then used to construct the annual earnings measure:

- Annualized earnings based on full-quarter status. A full-quarter worker is one who has positive earnings at a firm in the current (t), previous ($t-1$), and subsequent ($t+1$) quarters at the same firm. For workers who have worked at least one full quarter at a firm, 4 times the full-quarter average of earnings is used to construct the annual measure. Roughly 84% of the annual earnings are constructed this way.

- Annualized earnings based on continuous-quarter status. A continuous-quarter worker has positive earnings at a firm in the current (t) and previous ($t-1$) quarters or the current (t) and subsequent ($t+1$) quarters at the same firm. For workers who have not worked one full quarter during the calendar year, but who have at least one continuous quarter of employment, 8 times the average of continuous-quarter average earnings is used to construct the annual measure.[14]

- Annualized earnings based on reported quarterly earnings. For workers who are neither classified as full- or continuous-quarter for at least one quarter in the calendar year, 12 times the average of quarterly earnings is used to construct the annual measure.

For all observations, dummy variables are created to control for the type of annualized earnings measure that is used in the estimation of the wage model. The analysis sample is restricted to include employees who are between the ages of 25 and 65, have real annualized earnings between $1,000 and $1,000,000, and who work at firms with more than 10 employees that are not operating in agriculture or public administration Standard Industrial Classification (SIC) industry divisions.

Estimation of the multilevel model using data for all states providing data to the LEHD program is a computationally infeasible task. In order to estimate the model, I first select three states and then draw a

[14] The assumption is that a continuous-quarter worker (who is also not a full-quarter worker) has an expected employment duration of 0.50. Thus, observed continuous-quarter earnings are, in expectation, 50% of unobserved full-quarter earnings.

sample of workers from these states.[15] A requirement of the multilevel model is that workers are clustered within firms, thus, a procedure for sampling must guarantee that workers are clustered within firms. The sampling procedure developed in Woodcock (2003) is used and guarantees that the sample is representative of employment and that a minimum number of workers – here, at least 10 – are sampled from each firm. For 1997, a sample of firms is drawn with probabilities that are proportional to firm-level employment. Workers within those firms are sampled with probabilities that are inversely proportional to the firms' employment. Finally, the entire earnings histories for the sampled worker-firm pairs are included in the final dataset. The resulting datafile is a random sample of roughly 0.25% of workers in the three states and contains 283,507 annual observations on 55,267 individuals and 29,591 firms for the 1990-1998 time period.

Table 1 summarizes the list of variables used in the analysis, including a brief description of each. Tables 2a through 2c provide means and standard deviations of relevant characteristics for the analysis sample.

5 Estimation and Results

5.1 Specifications

I present estimates for four versions of the multilevel model. Each successive specification is a generalization of the preceding specification, reflecting the increasingly richer characterization of the firm-specific returns to human capital γ_{jt} and the variance matrix of the random error components in Γ. Specification 1 is a wage decomposition that allows for random person- and firm-intercepts and fixed returns to the observed components of human capital. This specification serves as the baseline wage model in which heterogeneity is captured only through person- and firm-specific intercepts. Specification 2 builds on Specification 1 by introducing the firm-specific returns to human capital; moreover, the variance matrix of these firm-specific returns is diagonal. In Specification 3, the assumption of a diagonal variance matrix of firm-specific returns is relaxed in favor of an unstructured one that permits covariation between the firm-specific returns and intercepts. Finally, Specification 4 retains the assumption of an unstructured variance matrix of firm-specific components and introduces observed firm characteristics – firm size and industrial classification – into the characterization of γ_{jt}. Appendix C relates each of the specifications to the model developed in Section 3. Characteristics of the four specifications are summarized in Table 3.

The four specifications are chosen for a number of reasons. Specification 1 is common in the literature on the decomposition of wages into unobserved, but identifiable, firm- and person-effects. Specification 2 permits an evaluation of the importance of allowing for firm-specific deviations from the economy-wide

[15] The identities of the states used in the analysis are witheld for confidentiality reasons. The three states chosen, however, are geographically dispersed across the United States and contain both large urban and rural areas.

returns to the elements of human capital and illustrates the relative importance of the estimated variance components. In Specification 3, the unstructured variance matrix of firm-specific returns permits an investigation of the significance of the covariation between the firm-specific returns and intercepts. Finally, Specification 4 introduces firm-level observed characteristics into the firm-specific returns in γ_{jt}. Estimation of this specification provides a direct assessment of the importance of both observed and unobserved firm characteristics in the determination of wages.

Specifications 1 through 3 are nested in terms of variance parameters and also share the same parameterization of the components for which coefficient estimates are fixed; for these specifications REML log-likelihood test statistics may be constructed to test which model best fits the data. Unfortunately, under the REML approach, there is no corresponding test statistic for comparing models where the fixed component of the model changes. Thus, Specification 4, which introduces firm-level covariates into the wage model cannot be tested against the preceding specifications. However, Specification 4 may be tested against specifications with the more restrictive variance structures – like those in Specifications 2 and 1 – that also retain the firm observed characteristics. Table 3 describes two additional specifications that are estimated to test directly the superiority of Specification 4.

Specification 5 is identical to 4 except for the variance matrix of the firm error components which is now diagonal (as in Specification 2); Specification 6 is similar to 5, though, the firm-specific returns are removed and only the firm- and person-specific intercepts are retained. Estimates of the variance parameters and coefficients for Specifications 5 and 6 are not discussed in this paper, though, the following section provides statistics pertaining to model fit to motivate a preference for Specification 4.

For all specifications, observed determinants of wages are fully interacted with sex so that fixed coefficient estimates on all person, demographic, and firm characteristics are reported separately for men and women. The firm-specific random returns to the components of human capital are also fully interacted with sex; thus, for example, separate variances (and covariances) of the dispersion in the returns to education are reported for men and women. Finally, the random firm-intercepts are not assumed to differ across men and women.[16]

5.2 Results

5.2.1 Model Selection

REML Likelihood Ratio Test Hypothesis tests involving the variance components of competing models estimated using REML techniques are usually conducted by constructing REML likelihood ratio tests (REMLRT) which are only valid in cases where the parameterization of the fixed coefficients is the same

[16] This is left to a future revision of the paper.

in both models. In the current analysis, Specification 1 is nested, in terms of variance parameters, within Specification 2, while Specification 2 is nested within Specification 3. In this sense, the REMLRT is useful tool for model selection.

The test statistic is constructed as

$$LLR = 2\log\left(\frac{l_{R2}}{l_{R1}}\right) = 2\left[\log\left(l_{R2}\right) - \log\left(l_{R1}\right)\right] \sim \chi^2_{r_2 - r_1}$$

where l_{R2} equals the log likelihood of the more general, unrestricted model; l_{R1} the log likelihood of the restricted model; r_2 is the number of estimated variance parameters in the unrestricted model; and r_1 is the number for the restricted model. The resulting test statistic is distributed χ^2 with $r_2 - r_1$ degrees of freedom.

Table 3 presents the estimated log likelihood and Akiake Information Criterion (AIC) for each of the four specifications. Calculations of the LLR suggest that Specification 2 is preferred over Specification 1 and Specification 3 over Specification 2 at all reported confidence levels. Thus, an unstructured variance structure seems to fit the data best. Recall that Specification 4 uses a variance matrix of random coefficients identical to that of Specification 3 but also includes the firm covariates. With these additional covariates, REMLRTs cannot be used to compare Specification 4 directly to Specifications 1 through 3. However, Specification 4 is preferred over Specification 5, which has the same variance parametrization as Specification 2; Specification 5 is preferred over Specification 6, which has the same variance parameterization as Specification 1. These results suggest that an fully unstructured variance matrix of firm returns is preferred, regardless of the choice of covariates to include in the fixed part of the model.

5.2.2 Coefficient and Variance Estimates

Recall that the estimation procedure first yields estimates of the model's variance parameters which are used to derive the fixed coefficient estimates – the BLUEs – and the random firm- and person-specific intercepts and firm-specific returns – the BLUPs. However, the estimates of the fixed coefficients are presented first as these are viewed, in the context of the firm-specific random returns, as the sample-wide average returns to the components of human capital. Thus, the discussion of the variance parameter estimates will be reviewed in light of these sample-wide averages. Predictors of the random components of the model are not reviewed in detail, though are used in the subsection that performs the analysis of variance decomposition.

Best Linear Unbiased Estimates (BLUEs) Table 4 presents the fixed coefficient estimates (BLUEs) for Specifications 1, 2, and 3 as well as the Ordinary Least Squares (OLS) estimates of these coefficients.

Recall that the data are demeaned prior to estimation, so no intercept (grand mean) is included in the estimations.

For Specification 1 (column 1), estimates of the fixed coefficients for men and women are reasonable in value and consistent with those found in the wage determination literature. For the quadratic in education, the coefficient on the first order term is positive for both men and women, though, higher for men than for women; the coefficient on the second order term is negative for both men and women, though larger for women. The size of these estimates implies an experience profile for men that is higher and more steeply sloped than the one for women. The return to education for men is positive and significant and implies a 4.67% increase in wages for each additional year of education; for women, the return to an additional year of education raises wages by 3.93%. The race dummy (non-white) is estimated to be negative for both men and women. Non-white men earn 36.55% less than white men; non-white women earn only 17.98% than white women.

The estimates for Specification 2 that includes the firm-specific returns to human capital and for Specification 4 that includes the firm-specific returns to human capital as well as an unstructured variance matrix for these returns are identical in sign and similar in magnitude. It is interesting, comparing the estimates from Specifications 1 through 3 to those emerging from OLS, that controlling for unobserved person- and firm-heterogeneity increases the wage penalty for non-white women (though not for men) and decreases the returns to education for both men and women.

Specification 4 includes the firm observed characteristics – firm employment and industry division of operation – as well as their interactions with the observed components of human capital. Table 5 presents the estimates of these fixed coefficients. Rows identify the characteristics of workers in the wage model; columns identify firm characteristics. SIC Division 1 (Mining and Construction) is the omitted industry group, thus, estimates under the "Constant" column heading refer to this group. Each row of estimates may be viewed as the effect of the observed firm characteristics on the wage impact of each observed component of human capital.

The first row suggests that for women, firm size and industry of operation have a significant effect on wages. For instance, the elasticity between firm size and real wages is 0.0531. The second and third rows suggest that the returns to experience are not significantly affected by firm size, are insignificant in Mining and Construction, and are significant across all other industry divisions. The return to education for women is not significantly affected by firm size, relatively small in Mining and Construction, and significantly higher in all other industry divisions (except for Transportation and Communications). Finally, the non-white wage penalty (25.77%) is higher than in Specifications 1 through 3 for all women except those working for firms in Professional Services.

For men, the elasticity between firm employment and wages is significant and roughly the same size as the estimate for women. And, the only significant industry division wage effects are in: Wholesale and Retail Trade; Finance, Insurance, and Real Estate; Services; and Professional Services. The first order experience component is slightly higher for larger firms. Otherwise, the experience profile is only significantly different for men in Finance, Insurance, and Real Estate. The return to education is slightly lower for men at larger firms but higher form men working in Finance, Insurance, and Real Estate and Professional Services. Finally, the non-white wage penalty is roughly the same size as in Specifications 1 through 3, but higher for men working in: Manufacturing; Wholesale and Retail Trade; Finance, Insurance, and Real Estate, and Services.

One issue is whether the results from Specification 4 are driven by the mixed effects specification, specifically the inclusion of the random returns and intercepts, or by the inclusion of firm covariates and their interactions with worker characteristics. One way to check the robustness of the specification is to estimate the fixed coefficients of the model using OLS (assuming no random person or firm effects). In terms of the pricing of the components of human capital, equation (2) becomes

$$\gamma_{jt} = g'_{jt}\beta^{(2)} \tag{12}$$

and the fixed coefficient estimates retain the same meaning as in Specification 4.

Table 5 presents the estimates for the OLS estimation. In comparing the estimates in Table 5 to those in Table 4, notice that many of the values change in magnitude, in sign, and in level of significance suggesting that the results are sensitive to the estimation technique. For instance, the OLS estimate of the elasticity of log of real earnings (wages) and firm size is -0.0823 for women and 0.1012 for men. The mixed model estimates in Specification 4 suggest an elasticity for both men and women that is just over 0.05.[17]

Overall, the fixed coefficient estimates emerging from Specifications 1 through 4 are consistent with the empirical wage determination literature. Men and women earn positive returns to education, have concave experience profiles, and earn a penalty to being non-white. Specification 4 suggests that large firms pay wage premia to both men and women and that industry of operation, even broadly defined, captures statistically significant variation in wages across firms. Finally, observed firm characteristics are shown to influence both average wages and, through their interaction with the observed components of human capital, are shown to influence a number of the returns to the observed components of human capital.

[17] The sensitivity of these estimates will be examined in more detail in the next revision of this paper.

Variance Parameter Estimates Table 7 presents estimates of the variance (and covariance) parameters for the Specifications 1 through 4. For both Specifications 1 and 2, the variance matrices are diagonal and estimated values are reported in rows. For Specifications 3 and 4, the variance matrices are unstructured in terms of the firm-specific returns (and intercept). The diagonal elements refer to the estimated variance of each term. Estimated covariances are reported below the diagonal and correlations are reported above the diagonal.

Recall that Specification 1 that includes only firm- and person-specific intercepts is treated as the baseline model as it is common in the wage decomposition literature. The estimates of the variance of the firm-specific intercept $\widehat{\sigma}^2_{\psi_1}$ (0.2606), the person-specific intercept $\widehat{\sigma}^2_{\alpha}$ (0.3215), and the residual $\widehat{\sigma}^2_{\varepsilon}$ (0.0682) are all statistically significant and have the following interpretation: that a one standard deviation increase in ψ_1, α, or ε increases wages by $\widehat{\sigma}_{\psi_1}$, $\widehat{\sigma}_{\alpha}$, or $\widehat{\sigma}_{\varepsilon}$ log points, respectively. The estimated variance of the person intercepts is larger than that of the firm intercepts, a common finding in the literature, suggesting that person-level heterogeneity is relatively more important than firm-level heterogeneity in wage variation.

The second row in Table 7 reports parameter estimates for Specification 2 that includes the firm-specific returns to human capital and assumes a diagonal variance matrix for these pay policy parameters. All estimated parameters are statistically significant. The introduction of the firm-specific returns lowers the estimated variance of the firm intercepts by over 30% from 0.2606 in Specification 1 to 0.1774 in Specification 2.[18] Both the estimated variance of the person intercepts and residual fall only slightly. These results suggest that the introduction of firm-specific returns to the components of human capital parses variation that would otherwise be explained by firm-intercepts alone. This is an important result. In the wage decomposition literature, firm-specific intercepts are often estimated and then, in a second step, decomposed into the portion explained by *firm-level* observed characteristics such as firm size or industry of employment.[19] Specification 2, on the other hand, ties firm-level unobserved, though identifiable, variation in wages across firms to observed *person-level* characteristics and, in this sense, explains a significant portion of the variation in wages across firms that would have otherwise been attributed to firm-intercepts alone.

Also reported for Specification 2 are the estimated variances of the firm-specific returns to human capital. The estimated variance of the returns to race (non-white) is large for both non-white men and women. The interpretation of the estimates is similar to the interpretation of those for the firm-intercept. For example, for a non-white woman, a one standard deviation increase in the firm-specific return to being non-white ψ_6 increases wages by $\widehat{\sigma}_{\psi_6}$ log points. Moreover, the total variance in wages due to the variance of average wages for non-whites across firms is given by $\widehat{\sigma}^2_{\psi_1} + \widehat{\sigma}^2_{\psi 2}$ for men and $\widehat{\sigma}^2_{\psi_1} + \widehat{\sigma}^2_{\psi 6}$ for women. The measure

[18] It would be interesting to introduce the additional firm parameters individually to assess how much of the decrease in $\widehat{\sigma}^2_{\psi_1}$ is associated with the addition of each component. This will be done in a future revision.

[19] Specification 4 essentially performs this decomposition by including firm-level observed characteristics in the wage model.

of race used in this analysis is relatively crude, so the estimated dispersion of wages across firms for this characteristic should be interpreted with caution. Dispersion in the firm-specific returns to wages for non-whites may reflect labor market discrimination – in the sense that firms have flexibility in setting wages for non-whites – though would also arise if non-white workers sort into a broad range of both high- and low-wage firms or certain high- and low-wage occupations within firms. Recall that the estimated sample-wide return (the BLUE) to non-white for both men and women implies that an average wage penalty is associated with this characteristic; dispersion in the returns to being non-white simply captures the dispersion in this penalty across firms. Another way to think about dispersion in returns to non-white is to consider what it would mean if these variance parameter estimates were statistically insignificant (from zero). In this case, the return race could still be negative, though, would be identical across firms.

The estimated variance of returns to education and experience for Specification 2 suggest that the dispersion in the returns to education across firms for women is slightly higher than for men ($\widehat{\sigma}^2_{\psi_3} > \widehat{\sigma}^2_{\psi_7}$) and that the earnings profiles for men are slightly more variable across firms than for women ($\widehat{\sigma}^2_{\psi_4} = \widehat{\sigma}^2_{\psi_5}$ and $\widehat{\sigma}^2_{\psi_8} > \widehat{\sigma}^2_{\psi_9}$). Dispersion in the returns to education and experience is interpreted in a manner that is slightly different than for the firm- and person-specific intercepts and return to race. These parameters measure the dispersion in the firm-specific prices of the components of human capital, thus the change in wages attributed to a change in the return to education for men, for example, is given by $dw_{ijt} = d\psi_3 * (education)$ and depends on a worker's level of education. For a man with 16 years of education, a one standard deviation increase in the return to the firm-specific education increases his real wage by 0.36 log points.

The assumption of a diagonal variance matrix of firm-specific pay components is relaxed in Specification 3. Table 7 presents variance and covariance estimates for this fully unstructured variance matrix. Diagonal elements contain the estimated variances of the pay policy parameters. Estimated covariances are reported below the diagonal and correlations are reported above the diagonal. The first thing to notice is the change in the estimated variance components between Specifications 2 and 3. The estimated variance of the firm-specific intercepts, the person-specific intercepts, and the residual fall only slightly. For both men and women, the estimated variances in the returns to race for both men and women increase as do the estimated variances of the return to education and the coefficients of the experience profiles. Thus, relaxing the structure of the variance matrix of pay policy parameters results in the attribution of more variation across firms in the returns to race, education, and experience. The estimation of covariance parameters permits a discussion of the extent to which firm pay policies are correlated within the sample. For instance, the first column of estimates for Specification 3 (below the first element of the diagonal) identify the covariation between the firm-specific intercepts and the firm-specific returns; the first row contains the correlations. The correlations between the firm-specific intercept and the returns to race for both men and women are

-0.0561 and -0.1571, respectively. Firms that tend to pay high average wages to all employees relative to other firms also tend to pay lower wages to non-whites relative to other firms. For both men and women, high average firm wages are positively correlated with the returns to education. For men, firm-specific intercepts are positively correlated with the return to the coefficient on the first order component of the experience profile and negatively correlated with the coefficient on the second order component suggesting that in high average wage firms, earnings profiles are relatively steeper for men for early years of experience but flatten in later years relative to the experience profiles of men at lower average wage firms. However, for women, firm intercepts are positively correlated with the coefficients on both terms of the experience profile suggesting that in firms that tend to pay high average wages, the return to experience is always greater, and increasingly so, over years of experience. Other correlations reveal interesting characteristics of firm compensation policy. The correlation between men and women in the returns to race (0.1728), education (0.5747), and the terms of the experience profiles (0.2227 on the first order term and 0.0555 on the second) suggest that firms pursue somewhat similar pay policies across gender: firms that tend to reward highly the characteristics of women also reward highly these characteristics for men.

Finally, Table 7 reports the variance parameter estimates for Specification 4 which includes the firm observed characteristics. The terms on the diagonal change only slightly relative to those reported for Specification 3. Interestingly, with the exception of the correlation between the firm-specific intercepts and the return to education for women, the correlations between the firm-specific intercepts and the other firm-specific returns increase in absolute value. And, with the exception of the correlation between the returns to the second order term in the experience profile for men and women, the correlations across sex in the specific returns to race, education, and the first order term of the quadratic in experience fall in absolute value.

The variance parameter estimates across the four specifications provide evidence that firms not only pay different wages to all of their workers but also pay different wages to certain types of workers. The estimated variances of the firm-specific returns to education, experience, and race are statistically significant and differ in size across men and women. The positive correlation across sex in the specific returns suggests a tendency for at least some firms to reward the human capital of both men and women similarly. Finally, firm average wages significantly co-vary with the returns to human capital. Non-whites tend to earn less at high average wage firms; education for both men and women are more highly rewarded in high average wage firms; and experience profiles become more steep, though exhibit turning points at lower levels of experience, in high average wage firms.

5.2.3 Correlations

Table 8 provides the correlations between the dependent variable, the total person effect, the total firm effect, and the observed time-varying covariates. The total person effect is composed of two parts: (1) the estimated unobserved person-level intercept $\widehat{\alpha}_i$ and (2) an observed component capturing the average returns to education, race, the race missing dummy, and the negative experience dummy. The total firm effect is also composed of two parts: (1) the unobserved firm-level intercept $\widehat{\psi}_1$ and (2) the firm-specific returns to the components of human capital. In the first line, for Specification 1, both the total person effect and the unobserved component $\widehat{\alpha}_i$ are more highly correlated with the log of real earnings than are the firm-specific intercepts. Moreover, correlation between the log of real earnings and the observed time varying covariates (last column) is lower than the correlations between earnings and any other component. These are standard findings in the wage decomposition literature. Notice that the total person effect and both the observed and unobserved components are slightly positively correlated. This finding is also consistent with the empirical literature and supports the notion that good workers sort into good firms. The existence of a slightly positive correlation between $\widehat{\alpha}$ and $\widehat{\psi}_1$ also suggests that the assumption of zero covariance between these terms (recall equation (11)) may be too restrictive.

The introduction of the firm-specific random returns in Specification 2 changes slightly the correlations between the log of real earnings and the model components in Specification 1. The firm-specific returns to human capital, however, are positively correlated with the log of real earnings (0.32); this correlation is larger than the correlation between the log of real earnings and the observed component of the person effect as well as the log of real earnings and the returns to the time varying covariates $X\beta$. The firm-specific returns exhibit a correlation with the unobserved person component $\widehat{\alpha}$ that is nearly zero, suggesting that workers with high innate ability do not sort into firms that pay high specific returns. However, the correlation between $\widehat{\alpha}$ and $\widehat{\psi}_1$ increases from 0.04 in Specification 1 to 0.09 in Specification 2 providing stronger evidence that high innate ability workers sort into firms that pay high average wages.

Correlations for Specification 3, where the variance matrix of firm-specific returns is unstructured, change only slightly, and unremarkably, relative to Specification 2. For Specification 4, however, the introduction of firm observed characteristics increases the positive correlation between the log of real earnings and the observed person component of wages. Moreover, the correlations between the log of real earnings and the firm intercepts falls, possibly reflecting the power of the firm-level covariates in explaining variation in wages that would have otherwise been captured by the firm-specific intercepts.

5.2.4 Analysis of Variance

The estimation approach – the use of REML to estimate the variance parameters and the Mixed Model Equations to predict and estimate the random and fixed coefficients – does not permit a straightforward decomposition of wages into variation attributable to the components of the wage model. For instance, there is no analog to a decomposition of variance using changes in model R-squared as in OLS. Even use of the estimated log likelihood function itself is of little value in assessing the relative importance of variation attributable to variables with estimated fixed coefficients and those with estimated random coefficients, as the value of the log likelihood itself is based only on the residual portion of the wage model. Nevertheless, this section attempts to assess the importance of the model's components in explaining variation in log wages through an analysis of variance exercise where the predicted random effects are treated as regressors in a linear regression on wages.

For each of the four specifications, a series of OLS regressions are estimated and include both the predicted random effects and the person-level covariates as regressors. Table 9 presents the results of this exercise.[20] For Specification 1, the baseline model $w_{ijt} = x_{ijt}^{(1)\prime}\beta^{(1)}$ is estimated yielding an R-squared of 0.1808. Introducing the predicted person intercept $\widehat{\alpha}_i$ increases the R-squared to 0.6785; adding the predicted firm intercept $\widehat{\psi}_1$ increases the R-squared to 0.9103; and, finally, including the estimated residuals $\widehat{\varepsilon}_{ijt}$ increases the R-squared to 1. Thus, observed worker characteristics (and control variables and time effects) explain roughly 18% of wage variation; predicted person intercepts explain roughly 50%; firm intercepts explain 23%; and the remaining 9% is explained by the residual.

For Specification 2, the firm-specific returns are included in the decomposition. The proportion of variation explained by the predicted person intercepts falls slightly to 46%; the proportion explained by the predicted firm intercepts falls to just under 20%; and the proportion explained by the residual falls slightly to 7%. Collectively, the firm-specific returns to human capital account for slightly under 9% of variation in wages: 4% is explained by the predicted returns to experience (for both men and women); 2.5% is explained by the predicted returns to education (for both men and women); and, 2.1% is explained by race (for both men and women). In Specification 3, the proportion of variation attributed to the firm-specific returns to human capital increases slightly and the proportion attributed to the residual falls slightly.

The observed firm characteristics are introduced in Specification 4 and explain a little over 5% of the variation in wages. The proportion of variation explained by the firm-specific returns to human capital is slightly over 9%; the proportion explained by firm intercepts alone is just over 17%. The relative size of these proportions underscores the importance of the firm-specific returns to human capital. The observed firm characteristics – firm size and industry division of operation – are broad measures in that they may not be precise enough to capture significant variation in wages across firms. For example, a larger sample of firms would permit the use of a finer measure of industry (e.g., 4-digit SIC) which would likely capture more variation in earnings. That the proportion of variation attributed to the firm-specific returns is greater than

[20] Results are relatively insensitive to the ordering of the components. Only one ordering is reported.

the proportion accounted for by observed firm characteristics probably overstates the relative unimportance of firm observed characteristics. However, including additional and more precise firm characteristics would simply result in a reallocation of explanatory power from the firm-specific intercepts to the set of observed characteristics. Collectively, observed firm characteristics and firm-specific intercepts account for 22% of the variation in wages. Given the 9% of variation attributed to the firm-specific returns to human capital, the proportion of wages explained by the firm-specific returns is roughly 30% of the total proportion of wages *explained by firms.*

6 Conclusion

That firms pay different wages is a known characteristic of labor markets. The reasons for these differentials, however, are not as well understood. While the results in this paper support previous findings of the existence of firm average wage differentials, this papers shows that not only one type of firm-specific pay applies to all workers. Men and women, whites and non-whites, and experienced and inexperienced workers earn different wages at different firms.

The observed firm characteristics included in this analysis suggest that large firms pay higher wages and that industry, even broadly defined, captures variation in wages across firms. However, these observed firm characteristics explain only a small portion of wage variation across firms. Including more precise measures of firm-level characteristics – ones that would influence the wages of all workers at the firm – may help explain why firms pay high average wages to all workers, though, would not necessarily decrease the importance of why firms pay specific wages to specific workers.

Statistically significant variation in firm-specific pay across a variety of worker characteristics suggests that compensation policy is specific to both the firm and the workers it employs. Although firm-specific returns may reflect the sorting of workers into particular types of firms or occupations within firms, they may also reflect the way firms differentially value the human capital that these characteristics measure. Firms may adopt production technologies or strategies that require a particular mix of skills. The empirical literature on firm wage differentials suggests that firms that pay high average wages may also be more productive firms. If highly productive technologies require highly skilled workers, then it seems reasonable for a compensation policy to include high average wages and high returns to education. The correlation between these pay policy parameters would then reflect the extent to which these production technologies exist. Finally, if monitoring is difficult in certain types of firms, firms may pay higher than average wages to all workers. If certain types of workers, those who are highly educated and experienced, are more difficult to monitor or sort into jobs that are more difficult to monitor, then firms may pay high returns to these characteristics. The existence of variation in the returns to education and experience alone may reflect monitoring problems associated with these characteristics of workers. A positive correlation between these specific returns and firm average wages may suggest that monitoring is difficult in particular firms and doubly so for the highly educated and experienced workers employed by those firms.

Moving deeper into the firm by specifying a model that allows unobserved firm characteristics to influence both average wages as well as worker characteristics, I find that approximately 30% of the variation in wages attributed to firms arises from the specific returns that firms pay to the components of human capital. These returns are significantly dispersed across employers, exhibit strong correlations across firms, and appear to be more important than observed firm characteristics in explaining variation in wages. The underlying compensation strategies that this model captures, however, is still an open question. Tying the policy measures estimated in this paper to other firm-level outcomes, for example, turnover and productivity, may help answer this question. This is a goal for future research.

Appendix A: REML Estimation of Variance Parameters

Transforming the Data

REML estimation of the variance components requires performing maximum likelihood estimation on the portion of the dependent variable – wages – net of the model's fixed effects. The discussion that follows is based on the one provided in (Searle et al. (1992)).

First, multiply (8) by some vector k yielding:

$$k'w = k'X\beta + k'X^{(2)}\Psi + k'D\alpha \tag{13}$$

where:

$$k'X\beta = 0 \ \forall \ \beta \tag{14}$$

Thus:

$$k'X = 0 \tag{15}$$

The vector k' is of a specific form, described further in Appendix M.4e of (Searle et al. (1992)). The number of linearly independent vectors k' is determined by the order $(N \times (kl + m))$ and rank r of the matrix of covariates in X for which fixed effects are estimated. These vectors are collected in a matrix K so that $K'X = 0$.

Estimation of Variance Components

From equation (8) we have:

$$w \sim N(X\beta, V) \tag{16}$$

where:

$$V = X^{(2)}(I_J \otimes \Gamma)X^{(2)\prime} + \sigma_\alpha^2 DD' + \sigma_\varepsilon^2 I_N \tag{17}$$

Pre-multiplying by K' yields:

$$K'w \sim N(0, K'VK) \tag{18}$$

and substituting into the log likelihood:

$$\log L_R = \frac{1}{2}(N-r)\log 2\pi - \frac{1}{2}\log|K'VK| - \frac{1}{2}w'K\left(K'VK\right)^{-1}K'w \quad (19)$$

Differentiating the likelihood with respect to the variance parameters σ^2_α, σ^2_ε, and the elements of Γ yields the first order conditions (REML equations) for the maximization of the likelihood[21].

$$\begin{aligned}
\frac{\partial L_R}{\partial \sigma^2_\alpha} &= -\frac{1}{2}trace\left[\left(K'VK\right)^{-1}K'DD'K\right] + \frac{1}{2}w'K\left(K'VK\right)^{-1}K'DD'K(K'VK)^{-1}K'w \\
\frac{\partial L_R}{\partial \sigma^2_\Gamma} &= -\frac{1}{2}trace\left[\left(K'VK\right)^{-1}K'X^{(2)}X^{(2)\prime}K\right] + \frac{1}{2}w'K\left(K'VK\right)^{-1}K'X^{(2)}X^{(2)\prime}K(K'VK)^{-1}K'w \\
\frac{\partial L_R}{\partial \sigma^2_\varepsilon} &= -\frac{1}{2}trace\left[\left(K'VK\right)^{-1}K'K\right] + \frac{1}{2}w'K\left(K'VK\right)^{-1}K'K(K'VK)^{-1}K'w
\end{aligned} \quad (20)$$

Setting the first order conditions equal to zero and using $P = K(K'VK)^{-1}K'$ yields:

$$\begin{aligned}
trace(PDD') &= w'PDD'Pw \\
trace(PX^{(2)}X^{(2)\prime}) &= w'PX^{(2)}X^{(2)\prime}Pw \\
trace(P) &= w'PPw
\end{aligned} \quad (21)$$

For the purpose of estimation, the conditions in (21) and the elements of the expected information matrix $\frac{\partial L_R}{\partial \sigma^2_i \partial \sigma^2_j}$ for $i, j = \alpha, \Gamma, \varepsilon$ are evaluated using the Average Information (AI) algorithm discussed in Gilmore, Thompson & Cullis (1995).

[21] $\frac{\partial L_R}{\partial \sigma^2_\Gamma}$ refers generally to the derivative of the log likelihood with respect to each of the elements contained in Γ.

Appendix B: Derivation of Best Linear Unbiased Estimates (BLUEs) and Best Linear Unbiased Predictors (BLUPs)

Using the estimates of $\widehat{\sigma}_{\alpha}^{2}$, $\widehat{\sigma}_{\varepsilon}^{2}$, and the elements of $\widehat{\Gamma}$ that emerge from REML estimation of equation (??), estimates of the fixed coefficients $\widehat{\beta}$ and the realized random coefficients $\widehat{\Psi}$ and $\widehat{\alpha}$ are obtained by maximizing the joint density of w, Ψ, and α with respect to β, Ψ, and α:

$$\max_{w.r.t.\ \beta,\Psi,\alpha} f(w,\Psi,\alpha) =$$

$$\frac{\exp\left\{-\frac{1}{2}\left[\left(w-X\beta-X^{(2)}\Psi-D\alpha\right)'\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]\left(w-X\beta-X^{(2)}\Psi-D\alpha\right)+\begin{bmatrix}\Psi\\ \alpha\end{bmatrix}\begin{bmatrix}I_J\otimes\Gamma & 0\\ 0 & \sigma_{\alpha}^{2}I_I\end{bmatrix}^{-1}\begin{bmatrix}\Psi & \alpha\end{bmatrix}\right]\right\}}{\left((2\pi)^{(N+q)}\left|[\sigma_{\varepsilon}^{2}I_N]\right|\left|\begin{bmatrix}I_J\otimes\Gamma & 0\\ 0 & \sigma_{\alpha}^{2}I_I\end{bmatrix}\right|\right)^{\frac{1}{2}}} \tag{22}$$

and equating the first order conditions to zero, yielding:

$$\begin{bmatrix} X'\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]X & X'\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]\begin{bmatrix}X^{(2)} & D\end{bmatrix} \\ \begin{bmatrix}X^{(2)\prime}\\ D'\end{bmatrix}\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]X & \begin{bmatrix}X^{(2)\prime}\\ D'\end{bmatrix}\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]\begin{bmatrix}X^{(2)} & D\end{bmatrix}+\begin{bmatrix}I_J\otimes\Gamma & 0\\ 0 & \sigma_{\alpha}^{2}I_I\end{bmatrix}^{-1}\end{bmatrix}\begin{bmatrix}\widehat{\beta}\\ \widehat{\Psi}\\ \widehat{\alpha}\end{bmatrix}$$

$$=\begin{bmatrix} X'\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]w \\ \begin{bmatrix}X^{(2)\prime}\\ D'\end{bmatrix}\left[\frac{1}{\sigma_{\varepsilon}^{2}}I_N\right]w\end{bmatrix} \tag{23}$$

the Mixed Model (or Henderson) Equations. Solving yields the Best Linear Unbiased Estimates (BLUEs) $\widehat{\beta}$, the vector of fixed coefficients, and the Best Linear Unbiased Predictors (BLUPs) $\widehat{\Psi}$, the random firm coefficients and intercepts, and $\widehat{\alpha}$, the random person intercepts:

$$\widehat{\beta}=BLUE\left[\beta\right]=\left(X'V^{-1}X\right)^{-1}X'V^{-1}w \tag{24}$$

$$\begin{bmatrix} \widehat{\Psi} \\ \widehat{\alpha} \end{bmatrix} = BLUP \begin{bmatrix} \Psi \\ \alpha \end{bmatrix} = \begin{bmatrix} \left(I_J \otimes \Gamma\right) X^{(2)\prime} & 0 \\ 0 & \sigma_\alpha^2 I_I D' \end{bmatrix} V^{-1} \left(w - X\widehat{\beta}\right) \tag{25}$$

where:

$$V = \sigma_\alpha^2 DD' + X^{(2)} \left(I_J \otimes \widehat{\Gamma}\right) X^{(2)\prime} + \sigma_\varepsilon^2 I_N \tag{26}$$

In practice, estimates of the variance parameters σ_α^2, σ_ε^2, and Γ emerging from REML estimation are used to derive $\widehat{\beta}$, $\widehat{\Psi}$, and $\widehat{\alpha}$. Estimates of the residuals are given by:

$$\widehat{\varepsilon} = w - X\widehat{\beta} - X^{(2)}\widehat{\Psi} - D\widehat{\alpha} \tag{27}$$

Equation (23) also illustrates two key aspects of mixed model approach in the context of prediction in a model with random effects. First, as $\left| \begin{bmatrix} I_J \otimes \Gamma & 0 \\ 0 & \sigma_\alpha^2 I_I \end{bmatrix}^{-1} \right| \to \infty$, the solutions for $\widehat{\beta}$, $\widehat{\Psi}$, and $\widehat{\alpha}$ tend to the generalized least squares estimates where $\widehat{\Psi}$ and $\widehat{\alpha}$ are treated as fixed instead of random. Given $\begin{bmatrix} I_J \otimes \Gamma & 0 \\ 0 & \sigma_\alpha^2 I_I \end{bmatrix}^{-1} < \infty$, the well known issues of identification that are present in models where $\widehat{\Psi}$ and $\widehat{\alpha}$ are treated as fixed are not present in the random effects case. Second, the off-diagonal elements of the left-hand-side of equation (23) reveal the impact of non-orthogonality assumption inherent in the modeling approach. These elements take into account the correlation of the designs of the person (D) and firm ($X^{(2)}$) effects with the covariates in X.

Appendix C: Details of Specifications

Specification 1

In Specification 1, the only firm-level random effect is an intercept. Thus, all covariates are included in $x_{ijt}^{(1)}$for which parameters are fixed. $x_{it}^{(2)}$ contains only a constant.

$$x_{ijt}^{(1)} = \begin{bmatrix} \text{male*(race missing dummy)} \\ \text{male*(negative experience dummy)} \\ \text{(1-male)*(race missing dummy)} \\ \text{(1-male)*(negative experience dummy)} \\ \text{male*(experience)} \\ \text{male*(experience)\^{}2/100} \\ \text{male*(education)} \\ \text{male*(nonwhite dummy)} \\ \text{(1-male)*(experience)} \\ \text{(1-male)*(experience)\^{}2/100} \\ \text{(1-male)*(education)} \\ \text{(1-male)*(nonwhite dummy)} \\ \text{time effect 1} \\ \vdots \\ \text{time effect 43} \end{bmatrix}, x_{it}^{(2)} = \begin{bmatrix} \text{constant} \end{bmatrix} \tag{28}$$

In terms of the general model, conditions (4) become:

$$f_j = \begin{bmatrix} f_{1j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ where } f_{1j} = 1, \Psi_j = \begin{bmatrix} \Psi_{1j} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \Gamma = \begin{bmatrix} \sigma^2_{\Gamma_{11}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{29}$$

and equation (7) is estimated and, in terms of notation, simplifies to:

$$w_{ijt} = x_{ijt}^{(1)\prime}\beta^{(1)} + x_{it}^{(2)\prime}\beta^{(2)} + \Psi_j + \alpha_i + \varepsilon_{ijt} \tag{30}$$

where Ψ_j contains only one element, an intercept.

Specification 2

In Specification 2, the returns to the components of (and controls for) human capital are permitted to vary across firms and are contained in the vector $x_{it}^{(2)}$. Sample-wide average returns are estimated for the components in $x_{ijt}^{(1)}$ as well.

$$x_{ijt}^{(1)} = \begin{bmatrix} \text{male*(race missing dummy)} \\ \text{male*(negative experience dummy)} \\ \text{(1-male)*(race missing dummy)} \\ \text{(1-male)*(negative experience dummy)} \\ \text{time effect 1} \\ \vdots \\ \text{time effect 43} \end{bmatrix}, x_{it}^{(2)} = \begin{bmatrix} \text{constant} \\ \text{male*(nonwhite dummy)} \\ \text{male*(education)} \\ \text{male*(experience)} \\ \text{male*(experience)\^{}2/100} \\ \text{(1-male)*(nonwhite dummy)} \\ \text{(1-male)*(education)} \\ \text{(1-male)*(experience)} \\ \text{(1-male)*(experience)\^{}2/100} \end{bmatrix} \tag{31}$$

In terms of the general model, conditions (4) become:

$$f_j = \begin{bmatrix} f_{1j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ where } f_{1j} = 1, \Psi_j = \begin{bmatrix} \Psi_{1j} \\ \vdots \\ \Psi_{lj} \end{bmatrix}, \Gamma = \begin{bmatrix} \sigma_{\Gamma_{11}}^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_{\Gamma_{ll}}^2 \end{bmatrix} \tag{32}$$

where the vector Ψ_j captures the firm-specific returns to the elements of $x_{it}^{(2)}$. Equation (7) is estimated.

Specification 3

Specification 3 extends Specification 2 by permitting Γ in (32) to be unstructured.

$$f_j = \begin{bmatrix} f_{1j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ where } f_{1j} = 1, \Psi_j = \begin{bmatrix} \Psi_{1j} \\ \vdots \\ \Psi_{lj} \end{bmatrix}, \Gamma = \begin{bmatrix} \sigma_{\Gamma_{11}}^2 & \cdots & \sigma_{\Gamma_{1l}}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{\Gamma_{l1}}^2 & \cdots & \sigma_{\Gamma_{ll}}^2 \end{bmatrix} \tag{33}$$

Equation (7) is estimated.

Specification 4

Finally, Specification 4 extends Specification 3 by modeling the firm-specific returns in equation (2) as a function of both the firm-specific random deviations in Ψ_j as well as observed firm-level employment and industry classification.

$$f_j = \begin{bmatrix} f_{1j} \\ \vdots \\ f_{lj} \end{bmatrix} \text{ where } f_{1j} = 1, \Psi_j = \begin{bmatrix} \Psi_{1j} \\ \vdots \\ \Psi_{lj} \end{bmatrix}, \Gamma = \begin{bmatrix} \sigma^2_{\Gamma_{11}} & \cdots & \sigma^2_{\Gamma_{1l}} \\ \vdots & \ddots & \vdots \\ \sigma^2_{\Gamma_{l1}} & \cdots & \sigma^2_{\Gamma_{ll}} \end{bmatrix} \tag{34}$$

Equation (7) is estimated.

References

Abowd, J., Kramarz, F. & Margolis, D. (1999). High wage workers and high wage firms, *Econometrica* **67**(2): 251–334.

Abowd, J. M., Kramarz, F. & Roux, S. (2005). Heterogeneity in frims' wages and mobility policies. Forthcoming in Structural Models of Wage and Employment Dynamics, Elsevier series on Contributions to Economic Analysis.

Abowd, J. M., Stephens, B., Vilhuber, L., Andersson, F., McKinney, K. L., Roemer, M. & Woodcock, S. (2005). The lehd infrastructure files and the creation of the quarterly workforce indicators. Prepared for the NBER Conference on Income and Wealth, Washington, DC. Forthcoming in conference volume.

Baker, G., Gibbs, M. & Holmstrom, B. (1995). The wage policy of a firm, *The Quarterly Journal of Economics* **107**: 921–957.

Bronars, S. G. & Famulari, M. (1997). Wage, tenure, and wage growth variation within and across establishments, *Journal of Labor Economics* **15**(2): 285–317.

Brown, C. & Medoff, J. (1989). The employer size-wage effect, *The Journal of Political Economy* **97**(5): 1027–1059.

Cappelli, P. & Neumark, D. (2001). Do "high performance" work practices improve establishment-level outcomes?, *Industrial Labor Relations Review* **54**(4): 737–775.

Cardoso, A. (1999). Firms' wage policies and the rise in labor market inequality: The case of portugal, *Industrial and Lobar Relations Review* **53**(1): 87–102.

Cardoso, A. (2000). Wage differentials across firms: An application of multilevel modelling, *Journal of Applied Econometrics* **15**(4): 343–354.

Davis, S. J. & Haltiwanger, J. (1995). Employer size and the wage structure in u.s. manufacturing. NBER Working Paper 5393.

Gilmore, A. R., Thompson, R. & Cullis, B. R. (1995). Average information reml: An efficient algorithm for variance parameter estimation in linear mixed models, *Biometrics* **51**(4): 1440–1450.

Goldstein, H. (1995). *Multilevel Statistical Models*, Kendall's Library of Statistics, Halsted Press.

Groshen, E. L. (1991). Sources of intra-industry wage dispersion: How much do employers matter?, *The Quarterly Journal of Economics* **106**(3): 869–884.

Groshen, E. & Levine, D. (1998). The rise and decline (?) of u.s. internal labor markets. Federal Reserve Bank of New York Research Paper 9819.

Ichniowski, C. & Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices, *Journal of Economic Perspectives* **17**(1): 155–180.

Lazear, E. (2003). Firm-specific human capital: A skill weights approach. NBER Working Paper 9679.

Lazear, E. P. (2000). The future of personnel economics, *The Economic Journal* **110**: 611–639.

McCulloch, C. E. & Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*, John Wiley and Sons, Inc.

Prendergast, C. (1999). The provision of incentives in firms, *Journal of Economic Literature* **37**(1): 7–63.

Raudenbush, S. & Bryk, A. (1986). A hierarchical model for studying school effects, *Sociology of Education* pp. 1–17.

Searle, S. R., Casella, G. & McCulloch, C. E. (1992). *Variance Components*, John Wiley and Sons, Inc.

Woodcock, S. D. (2003). Agent heterogeneity and learning: An application to labor markets. LEHD Technical Paper TP-2002-20.

Table 1: Variables in Anaylsis

Demographics	
Education	Based on statistical match to Decennial Census 1990
Non-white	Based on race variable in Census NUMIDENT
Race Missing	Based on race variable in Census NUMIDENT
Sex	Based on sex variable in Census NUMIDENT
Job Characteristics	
ln(Annualized Real Wage)	Annualized wage measure based on quarterly earnings (UI)
Experience	In the first quarter that an individual appears in the data, potential experience is calculated as age at the beginning of the quarter minus years of education minus 6; potential experience is then augmented with each additional year of observed experience (years of positive annualized earnings).
Age	Base on date-of-birth measure reported in the Census NUMIDENT
Negative Experience Dummy	Equals 1 if Experience is calculated negative
Firm Characteristics	
ln(Firm Employment)	Natural log of the sum of workers with positive annualized earnings
SIC 1	Industry: Mining and Construction
SIC 2	Industry: Manufacturing
SIC 3	Industry: Manufacturing
SIC 4	Industry: Transporation, Communications, Electric, Gas, and Sanitary Services
SIC 5	Industry: Wholesale and Retail Trade
SIC 6	Industry: Finance, Insurance, and Real Estate
SIC 7	Industry: Services
SIC 8	Industry: Professional Services
Time Dummies	
4 Full Quarters Worked 1990	Dummy
4 Full Quarters Worked 1991	Dummy
4 Full Quarters Worked 1992	Dummy
4 Full Quarters Worked 1993	Dummy
4 Full Quarters Worked 1994	Dummy
4 Full Quarters Worked 1995	Dummy
4 Full Quarters Worked 1996	Dummy
4 Full Quarters Worked 1997	Dummy
4 Full Quarters Worked 1998	Dummy
Less Than 4 Full Quarters Worked 1990	Dummy
Less Than 4 Full Quarters Worked 1991	Dummy
Less Than 4 Full Quarters Worked 1992	Dummy
Less Than 4 Full Quarters Worked 1993	Dummy
Less Than 4 Full Quarters Worked 1994	Dummy
Less Than 4 Full Quarters Worked 1995	Dummy
Less Than 4 Full Quarters Worked 1996	Dummy
Less Than 4 Full Quarters Worked 1997	Dummy
Less Than 4 Full Quarters Worked 1998	Dummy
Discontinuous Employment Dummy	Dummy
0 Full Quarters Worked	Dummy
1 Full Quarters Worked	Dummy
2 Full Quarters Worked	Dummy
3 Full Quarters Worked	Dummy
4 Full Quarters Worked	Dummy

Table 2a: Sample Demographics

	Men		Women	
	Analysis Sample		Analysis Sample	
	Mean	*Standard Deviation*	Mean	*Standard Deviation*
Proportion Male	0.5290	0.4992		
Education	12.5484	2.9799	12.7728	2.5678
Proportion Non-white	0.3216	0.4671	0.3335	0.4715
Proportion Race Missing	0.0422	0.2011	0.0321	0.1763

Table 2b: Sample Job Characteristics

	Men		Women	
	Analysis Sample		Analysis Sample	
	Mean	*Standard Deviation*	Mean	*Standard Deviation*
ln(Annualized Real Wage)	10.6035	0.7441	10.2452	0.6706
Experience	21.9814	9.9102	21.9719	9.8157
Age	40.1197	9.5078	40.2895	9.3661
ln(Firm Employment)	6.1190	2.3237	6.5018	2.2787
Proportion in One-Digit SIC 1	0.0816	0.2738	0.0210	0.1434
Proportion in One-Digit SIC 2	0.0782	0.2685	0.0608	0.2390
Proportion in One-Digit SIC 3	0.1799	0.3841	0.0906	0.2870
Proportion in One-Digit SIC 4	0.1067	0.3087	0.0676	0.2511
Proportion in One-Digit SIC 5	0.2125	0.4091	0.1752	0.3802
Proportion in One-Digit SIC 6	0.0614	0.2401	0.1062	0.3081
Proportion in One-Digit SIC 7	0.1091	0.3118	0.0880	0.2833
Proportion in One-Digit SIC 8	0.1706	0.3761	0.3906	0.4879
4 Full Quarters Worked 1990	0.0461	0.2096	0.0478	0.2134
4 Full Quarters Worked 1991	0.0483	0.2144	0.0498	0.2176
4 Full Quarters Worked 1992	0.0623	0.2418	0.0629	0.2429
4 Full Quarters Worked 1993	0.0728	0.2598	0.0756	0.2643
4 Full Quarters Worked 1994	0.0897	0.2857	0.0915	0.2883
4 Full Quarters Worked 1995	0.1020	0.3027	0.1047	0.3062
4 Full Quarters Worked 1996	0.1252	0.3309	0.1295	0.3358
4 Full Quarters Worked 1997	0.1363	0.3431	0.1392	0.3462
4 Full Quarters Worked 1998	0.1257	0.3315	0.1292	0.3354
Less Than 4 Full Quarters Worked 1990	0.0122	0.1098	0.0116	0.1070
Less Than 4 Full Quarters Worked 1991	0.0125	0.1112	0.0123	0.1101
Less Than 4 Full Quarters Worked 1992	0.0158	0.1246	0.0163	0.1265
Less Than 4 Full Quarters Worked 1993	0.0179	0.1326	0.0168	0.1287
Less Than 4 Full Quarters Worked 1994	0.0195	0.1381	0.0177	0.1317
Less Than 4 Full Quarters Worked 1995	0.0273	0.1631	0.0246	0.1549
Less Than 4 Full Quarters Worked 1996	0.0376	0.1903	0.0345	0.1824
Less Than 4 Full Quarters Worked 1997	0.0808	0.2725	0.0714	0.2574
Less Than 4 Full Quarters Worked 1998	0.0408	0.1979	0.0388	0.1932
Discontinuous Employment Dummy	0.0122	0.1099	0.0096	0.0976
0 Full Quarters Worked	0.0528	0.2237	0.0437	0.2043
1 Full Quarters Worked	0.0630	0.2429	0.0592	0.2360
2 Full Quarters Worked	0.0679	0.2516	0.0637	0.2443
3 Full Quarters Worked	0.0656	0.2476	0.0628	0.2425
4 Full Quarters Worked	0.7507	0.4326	0.7707	0.4204

Table 2c: Sample Job Characteristics and Firm Interactions

	Men		Women	
	Analysis Sample		Analysis Sample	
	Mean	*Standard Deviation*	Mean	*Standard Deviation*
Education Interacted with ln(Firm Employment)	77.8903	35.5628	83.9931	34.4374
Education Interacted with One-Digit SIC 1	0.9723	3.3710	0.2634	1.8376
Education Interacted with One-Digit SIC 2	0.9511	3.3719	0.7315	2.9578
Education Interacted with One-Digit SIC 3	2.2268	4.9013	1.1090	3.6003
Education Interacted with One-Digit SIC 4	1.3239	3.9529	0.8601	3.2610
Education Interacted with One-Digit SIC 5	2.6375	5.2537	2.1731	4.8323
Education Interacted with One-Digit SIC 6	0.8346	3.3367	1.3848	4.0962
Education Interacted with One-Digit SIC 7	1.4226	4.1722	1.1401	3.7530
Education Interacted with One-Digit SIC 8	2.3280	5.2592	5.2352	6.7051
Non-white Interacted with ln(Firm Employment)	1.7608	3.0250	2.1190	3.3776
Non-white Interacted with One-Digit SIC 1	0.0176	0.1316	0.0057	0.0752
Non-white Interacted with One-Digit SIC 2	0.0275	0.1635	0.0270	0.1620
Non-white Interacted with One-Digit SIC 3	0.0552	0.2284	0.0352	0.1844
Non-white Interacted with One-Digit SIC 4	0.0279	0.1647	0.0231	0.1502
Non-white Interacted with One-Digit SIC 5	0.0639	0.2446	0.0511	0.2202
Non-white Interacted with One-Digit SIC 6	0.0148	0.1206	0.0311	0.1736
Non-white Interacted with One-Digit SIC 7	0.0376	0.1903	0.0319	0.1756
Non-white Interacted with One-Digit SIC 8	0.0433	0.2036	0.1105	0.3135
Experience Interacted with ln(Firm Employment)	135.7960	84.6476	143.5200	84.9780
Experience Interacted with One-Digit SIC 1	1.8232	6.7139	0.4793	3.5697
Experience Interacted with One-Digit SIC 2	1.7600	6.6453	1.3732	5.9306
Experience Interacted with One-Digit SIC 3	4.1983	9.9064	2.0685	7.1471
Experience Interacted with One-Digit SIC 4	2.4353	7.7308	1.4933	6.0308
Experience Interacted with One-Digit SIC 5	4.4610	9.6892	3.6959	9.0445
Experience Interacted with One-Digit SIC 6	1.2260	5.3552	2.1972	7.1314
Experience Interacted with One-Digit SIC 7	2.2161	7.1546	1.7777	6.4404
Experience Interacted with One-Digit SIC 8	3.8614	9.4503	8.8867	12.6636

Table 3: Model Selection

	Specification 1	Specification 2	Specification 3	**Specification 4**	Specification 5	Specification 6
Random Intercepts	YES	YES	YES	**YES**	YES	YES
Random Slopes	NO	YES	YES	**YES**	YES	NO
Variance Structure	DIAGONAL	DIAGONAL	UNSTRUCTURED	**UNSTRUCTURED**	DIAGONAL	DIAGONAL
Firm Observed Characteristics	NO	NO	NO	**YES**	YES	YES
Log Likelihood	125600	132540	135001	**135834**	133335	126488
AIC	-251196	-265060	-269910	**-271576**	-266650	-252972
Test of Model Preference		2 over 1	3 over 2	**4 over 5**	5 over 6	
REML Likelihood Ratio		13880	4922	**4998**	13694	
d.o.f.		8	36	**36**	8	

Table 4: Fixed Coefficient Estimates for Specifications 1 - 3

			Specification 1	Specification 2	Specification 3	OLS
Female	Experience	Estimate	0.0293 *	0.0288 *	0.0296 *	0.0273 *
		Standard Error	0.0009	0.0010	0.0013	0.0009
	$Experience^2/100$	Estimate	-0.0462 *	-0.0429 *	-0.0457 *	-0.0534 *
		Standard Error	0.0019	0.0021	0.0027	0.0018
	Education	Estimate	0.0393 *	0.0414 *	0.0450 *	0.0743 *
		Standard Error	0.0015	0.0016	0.0018	0.0008
	Non-white Dummy	Estimate	-0.1798 *	-0.1827 *	-0.1917 *	-0.1387 *
		Standard Error	0.0083	0.0099	0.0100	0.0041
	Race Missing Dummy	Estimate	-0.1684 *	-0.1791 *	-0.1858 *	0.0758 *
		Standard Error	0.0216	0.0224	0.0225	0.0131
	Negative Experience Dummy	Estimate	-0.0562 *	-0.0632 *	-0.0703 *	-0.0390 *
		Standard Error	0.0280	0.0286	0.0293	0.0082
Male	Experience	Estimate	0.0522 *	0.0513 *	0.0544 *	0.0423 *
		Standard Error	0.0009	0.0009	0.0013	0.0008
	$Experience^2/100$	Estimate	-0.0826 *	-0.0782 *	-0.0841 *	-0.0720 *
		Standard Error	0.0017	0.0020	0.0027	0.0017
	Education	Estimate	0.0467 *	0.0490 *	0.0525 *	0.0632 *
		Standard Error	0.0012	0.0013	0.0015	0.0007
	Non-white Dummy	Estimate	-0.3655 *	-0.3724 *	-0.3824 *	-0.3695 *
		Standard Error	0.0079	0.0094	0.0095	0.0039
	Race Missing Dummy	Estimate	-0.2163 *	-0.2122 *	-0.2060 *	0.0002
		Standard Error	0.0179	0.0182	0.0184	0.0093
	Negative Experience Dummy	Estimate	-0.0485	-0.0512	-0.0475	-0.0953 *
		Standard Error	0.0288	0.0295	0.0303	0.0083

** indicates significance at 5%*

Table 5: Fixed Coefficient Estimates for Specification 4

			Constant	ln(Firm Employment)	Manufacturing (SIC 2)	Manufacturing (SIC 3)	Transporation, Communications, Electric, Gas, and Sanitary Services (SIC 4)	Wholesale and Retail Trade (SIC 5)	Finance, Insurance, and Real Estate (SIC 6)	Services (SIC 7)	Professional Services (SIC 8)
Female	Constant	Estimate		0.0531 *	-0.5142 *	-0.5810 *	-0.4411 *	-0.7146 *	-0.4874 *	-0.7110 *	-0.7797 *
		Standard Error		0.0098	0.1154	0.1071	0.1210	0.0973	0.1120	0.1021	0.0941
	Experience	Estimate	-0.0016	0.0006	0.0193 *	0.0252 *	0.0301 *	0.0290 *	0.0223 *	0.0318 *	0.0278 *
		Standard Error	0.0068	0.0006	0.0075	0.0071	0.0079	0.0067	0.0071	0.0068	0.0065
	Experience^2/100	Estimate	0.0030	-0.0001	-0.0374 *	-0.0477 *	-0.0592 *	-0.0543 *	-0.0393 *	-0.0642 *	-0.0485 *
		Standard Error	0.0147	0.0012	0.0159	0.0150	0.0168	0.0142	0.0153	0.0146	0.0138
	Education	Estimate	0.0273 *	-0.0009	0.0170 *	0.0174 *	0.0075	0.0179 *	0.0190 *	0.0181 *	0.0297 *
		Standard Error	0.0062	0.0006	0.0070	0.0064	0.0071	0.0060	0.0070	0.0063	0.0058
	Non-white Dummy	Estimate	-0.2577 *	0.0019	-0.0455	0.0278	0.0600	0.0457	0.0818	0.0397	0.0835 *
		Standard Error	0.0432	0.0035	0.0445	0.0423	0.0465	0.0403	0.0437	0.0407	0.0393
	Race Missing Dummy	Estimate	-0.1734 *								
		Standard Error	0.0223								
	Negative Experience Dummy	Estimate	-0.0569								
		Standard Error	0.0290								
Male	Constant	Estimate		0.0516 *	-0.0764	-0.0849	0.1043	-0.1742 *	-0.2282 *	-0.1643 *	-0.2710 *
		Standard Error		0.0089	0.0796	0.0714	0.0809	0.0648	0.0918	0.0682	0.0673
	Experience	Estimate	0.0394 *	0.0014 *	-0.0052	0.0018	-0.0034	0.0075	0.0147 *	0.0067	0.0074
		Standard Error	0.0043	0.0006	0.0048	0.0044	0.0051	0.0040	0.0055	0.0042	0.0043
	Experience^2/100	Estimate	-0.0643 *	-0.0017	0.0093	-0.0001	0.0034	-0.0130	-0.0290 *	-0.0180	-0.0144
		Standard Error	0.0094	0.0012	0.0102	0.0093	0.0110	0.0088	0.0119	0.0092	0.0091
	Education	Estimate	0.0509 *	-0.0010 *	0.0079	0.0020	-0.0052	0.0020	0.0138 *	0.0026	0.0124 *
		Standard Error	0.0039	0.0005	0.0044	0.0040	0.0044	0.0036	0.0052	0.0037	0.0037
	Non-white Dummy	Estimate	-0.3685 *	0.0064	-0.0622 *	-0.0142	-0.0521	-0.0581 *	-0.0843 *	-0.0974 *	-0.0288
		Standard Error	0.0267	0.0033	0.0286	0.0259	0.0298	0.0241	0.0343	0.0250	0.0256
	Race Missing Dummy	Estimate	-0.1999 *								
		Standard Error	0.0182								
	Negative Experience Dummy	Estimate	-0.0362								
		Standard Error	0.0300								

** indicates significance at 5%*

Table 6: OLS Coefficient Estimates Corresponding to Specification 4

			Constant	ln(Firm Employment)	Manufacturing (SIC 2)	Manufacturing (SIC 3)	Transporation, Communications, Electric, Gas, and Sanitary Services (SIC 4)	Wholesale and Retail Trade (SIC 5)	Finance, Insurance, and Real Estate (SIC 6)	Services (SIC 7)	Professional Services (SIC 8)
Female	Constant	Estimate		-0.0823 *	0.2363	0.2427	0.6483 *	-0.0754	0.1038	-0.0550	-0.0477
		Standard Error		0.0067	0.1243	0.1307	0.1226	0.1107	0.1162	0.1144	0.1090
	Experience	Estimate	-0.0022	0.0039 *	-0.0002	0.0062	0.0084	0.0024	0.0034	0.0021	-0.0033
		Standard Error	0.0061	0.0004	0.0068	0.0073	0.0069	0.0061	0.0064	0.0063	0.0060
	Experience^2/100	Estimate	0.0035	-0.0060 *	-0.0099	-0.0193	-0.0298	-0.0172	-0.0235	-0.0233	-0.0019
		Standard Error	0.0132	0.0008	0.0147	0.0156	0.0150	0.0132	0.0140	0.0138	0.0131
	Education	Estimate	0.0535 *	0.0037 *	-0.0258 *	-0.0236 *	-0.0466 *	-0.0178 *	-0.0055	-0.0065	-0.0027
		Standard Error	0.0064	0.0004	0.0072	0.0074	0.0070	0.0064	0.0067	0.0066	0.0063
	Non-white Dummy	Estimate	-0.0665	0.0119 *	-0.2724 *	-0.0899 *	-0.0181	-0.1872 *	-0.1835 *	-0.3398 *	-0.1280 *
		Standard Error	0.0374	0.0021	0.0401	0.0415	0.0395	0.0373	0.0388	0.0379	0.0367
	Race Missing Dummy	Estimate	0.0702 *								
		Standard Error	0.0128								
	Negative Experience Dummy	Estimate	-0.0459 *								
		Standard Error	0.0080								
Male	Constant	Estimate		0.1012 *	-0.1321 *	-0.2843 *	-0.1366 *	-0.5127 *	-0.5162 *	-0.4277 *	-0.6590 *
		Standard Error		0.0064	0.0575	0.0543	0.0517	0.0429	0.0617	0.0496	0.0461
	Experience	Estimate	0.0323 *	-0.0003	0.0026	0.0044	0.0140 *	0.0200 *	0.0161 *	0.0040	0.0126 *
		Standard Error	0.0028	0.0004	0.0036	0.0034	0.0033	0.0027	0.0040	0.0031	0.0029
	Experience^2/100	Estimate	-0.0542 *	0.0019 *	-0.0113	-0.0093	-0.0347 *	-0.0451 *	-0.0407 *	-0.0256 *	-0.0355 *
		Standard Error	0.0059	0.0008	0.0075	0.0070	0.0069	0.0057	0.0086	0.0067	0.0062
	Education	Estimate	0.0807 *	-0.0074 *	0.0086 *	0.0249 *	0.0130 *	0.0194 *	0.0453 *	0.0304 *	0.0438 *
		Standard Error	0.0024	0.0003	0.0032	0.0030	0.0029	0.0024	0.0034	0.0027	0.0026
	Non-white Dummy	Estimate	-0.4264 *	0.0150 *	-0.0223	0.1624 *	0.0743 *	-0.0275	-0.2963 *	-0.1730 *	-0.0478 *
		Standard Error	0.0152	0.0021	0.0189	0.0184	0.0169	0.0146	0.0224	0.0164	0.0158
	Race Missing Dummy	Estimate	-0.0044								
		Standard Error	0.0091								
	Negative Experience Dummy	Estimate	-0.0894 *								
		Standard Error	0.0081								

** indicates significance at 5%*

Table 7: Variance Estimates

			Firm Intercept ψ_1	Male: Non-white ψ_2	Male: Education ψ_3	Male: Experience ψ_4	Male: $\frac{\text{Experience}^2}{100}$ ψ_5	Female: Non-white ψ_6	Female: Education ψ_7	Female: Experience ψ_8	Female: $\frac{\text{Experience}^2}{100}$ ψ_9	Person Intercept α	Residual ε
Specification 1			**0.2606** *									**0.3215** *	**0.0682** *
Specification 2			**0.1774** *	**0.1283** *	**0.0005** *	**0.0001** *	**0.0010** *	**0.0940** *	**0.0004** *	**0.0001** *	**0.0006** *	**0.3080** *	**0.0579** *
Specification 3													
	Firm Intercept	ψ_1	**0.1755** *	-0.0561	0.0683	0.1268	-0.0948	-0.1571	0.0834	0.0836	-0.0799		
Male	Non-white	ψ_2	-0.0091 *	**0.1502** *	-0.0314	0.0456	-0.0462	0.1728	0.1176	0.1226	-0.0875		
	Education	ψ_3	0.0015 *	-0.0006	**0.0026** *	0.0044	0.1146	0.1084	0.5747	0.3447	-0.2321		
	Experience	ψ_4	0.0033 *	0.0011	0.0000	**0.0040** *	-0.9534	0.1236	0.4836	0.2227	-0.1132		
	Experience^2 /100	ψ_5	-0.0052 *	-0.0023	0.0008 *	-0.0079 *	**0.0172** *	-0.1122	-0.3490	-0.1332	0.0555		
Female	Non-white	ψ_6	-0.0245 *	0.0250 *	0.0021 *	0.0029 *	-0.0055 *	**0.1391** *	0.0240	-0.0811	0.1316		
	Education	ψ_7	0.0020 *	0.0026 *	0.0017 *	0.0017 *	-0.0026 *	0.0005	**0.0033** *	-0.0621	0.1443		
	Experience	ψ_8	0.0018 *	0.0025 *	0.0009 *	0.0007 *	-0.0009 *	-0.0016 *	-0.0002	**0.0027** *	-0.9508		
	Experience^2 /100	ψ_9	-0.0036 *	-0.0036	-0.0013 *	-0.0008 *	0.0008	0.0052 *	0.0009 *	-0.0053 *	**0.0114** *		
Person Intercept		α										**0.3022** *	
Residual		ε											**0.0551** *
Specification 4													
	Firm Intercept	ψ_1	**0.1618** *	-0.0869	0.0736	0.1368	-0.1165	-0.1601	0.0696	0.0969	-0.1060		
Male	Non-white	ψ_2	-0.0135 *	**0.1498** *	-0.0283	0.0340	-0.0341	0.1693	0.1158	0.1202	-0.0884		
	Education	ψ_3	0.0015 *	-0.0006	**0.0027** *	-0.0051	0.1167	0.1154	0.5704	0.3456	-0.2393		
	Experience	ψ_4	0.0034 *	0.0008	0.0000	**0.0038** *	-0.9542	0.1164	0.4729	0.2160	-0.1157		
	Experience^2 /100	ψ_5	-0.0061 *	-0.0017	0.0008 *	-0.0076 *	**0.0167** *	-0.1033	-0.3437	-0.1284	0.0570		
Female	Non-white	ψ_6	-0.0240 *	0.0244 *	0.0022 *	0.0027 *	-0.0050 *	**0.1389** *	0.0254	-0.0826	0.1314		
	Education	ψ_7	0.0016 *	0.0026 *	0.0017 *	0.0017 *	-0.0025 *	0.0005	**0.0033** *	-0.0714	0.1415		
	Experience	ψ_8	0.0020 *	0.0024 *	0.0009 *	0.0007 *	-0.0009 *	-0.0016 *	-0.0002 *	**0.0027** *	-0.9530		
	Experience^2 /100	ψ_9	-0.0046 *	-0.0037 *	-0.0013 *	-0.0008 *	0.0008	0.0052 *	0.0009 *	-0.0053 *	**0.0115** *		
Person Intercept		α										**0.2946** *	
Residual		ε											**0.0552** *

** indicates significance at 5%*

For Specifications 1 and 2, the rows correspond to estimated variances of each parameter indicated in the column heading. For Specifications 3 and 4, variance estimates are on the diagonal, covariance estimates are below the diagonal, and correlations are above the diagonal.

Table 8: Correlations

		ln(Annualized Real Earnings)	Total Person Effect	Unobserved Component (α)	Observed Component[a]	Total Firm Effect ($X^{(2)}\psi$)	Firm Intercept (ψ_1)	Firm-Specific Return[b]	Time-Varying Covariates ($X\beta$)
Specification 1	ln(Annualized Real Earnings)	1	0.76	0.70	0.29	n.a.	0.55	n.a.	0.26
	Total Person Effect		1	0.94	0.34	n.a.	0.07	n.a.	-0.05
	Unobserved Component (α)			1	-0.01	n.a.	0.04	n.a.	-0.06
	Observed Component[a]				1	n.a.	0.08	n.a.	0.01
	Firm Intercept (ψ_1)						1	n.a.	0.04
	Time-Varying Covariates ($X\beta$)								1
Specification 2	ln(Annualized Real Earnings)	1	0.74	0.68	0.29	0.61	0.54	0.32	0.26
	Total Person Effect		1	0.92	0.37	0.09	0.12	-0.01	-0.06
	Unobserved Component (α)			1	-0.01	0.07	0.09	-0.01	-0.07
	Observed Component[a]				1	0.06	0.08	-0.02	0.01
	Total Firm Effect ($X^{(2)}\psi$)					1	0.85	0.59	0.02
	Firm Intercept (ψ_1)						1	0.08	0.03
	Firm-Specific Return[b]							1	-0.01
	Time-Varying Covariates ($X\beta$)								1
Specification 3	ln(Annualized Real Earnings)	1	0.74	0.67	0.28	0.62	0.55	0.33	0.26
	Total Person Effect		1	0.91	0.40	0.10	0.13	-0.01	-0.06
	Unobserved Component (α)			1	-0.01	0.09	0.10	0.02	-0.07
	Observed Component[a]				1	0.03	0.08	-0.06	0.00
	Total Firm Effect ($X^{(2)}\psi$)					1	0.84	0.60	0.01
	Firm Intercept (ψ_1)						1	0.08	0.03
	Firm-Specific Return[b]							1	-0.03
	Time-Varying Covariates ($X\beta$)								1
Specification 4									
	ln(Annualized Real Earnings)	1	0.76	0.67	0.35	0.57	0.50	0.33	0.28
	Total Person Effect		1	0.89	0.44	0.09	0.12	-0.01	0.01
	Unobserved Component (α)			1	-0.02	0.07	0.08	0.02	0.00
	Observed Component[a]				1	0.06	0.11	-0.05	0.03
	Total Firm Effect ($X^{(2)}\psi$)					1	0.83	0.61	-0.08
	Firm Intercept (ψ_1)						1	0.08	-0.08
	Firm-Specific Return[b]							1	-0.03
	Time-Varying Covariates ($X\beta$)								1

[a] The observed component is the portion of wages attributed to: education, non-white (dummy), race missing (dummy), and negative experience (dummy).

[b] The firm-specific return is the portion of wages attributed to the firm-specific returns to person-level characteristics.

Table 9: Analysis of Variance

	Worker Covariates and Controls	Firm Covariates and Interactions	Predicted Person Intercept	Predicted Firm Intercept	Predicted Experience	Predicted Education	Predicted Non-White	Estimated Residual
Specification 1								
R-squared	0.1808	n.a.	0.6785	0.9103	n.a.	n.a.	n.a.	1.0000
Δ R-squared			0.4977	0.2318				0.0897
Specification 2								
R-squared	0.1808	n.a.	0.6432	0.8403	0.8810	0.9063	0.9273	1.0000
Δ R-squared			0.4624	0.1971	0.0407	0.0253	0.0210	0.0727
Specification 3								
R-squared	0.1808	n.a.	0.6404	0.8379	0.8681	0.9078	0.9323	1.0000
Δ R-squared			0.4596	0.1975	0.0302	0.0397	0.0245	0.0677
Specification 4								
R-squared	0.1808	0.2325	0.6671	0.8390	0.8679	0.9076	0.9322	1.0000
Δ R-squared		0.0517	0.4346	0.1719	0.0289	0.0397	0.0246	0.0678