Econometrica, Vol. 67, No. 2 (March, 1999), 251–333

HIGH WAGE WORKERS AND HIGH WAGE FIRMS

BY JOHN M. ABOWD, FRANCIS KRAMARZ, AND DAVID N. MARGOLIS[1]

We study a longitudinal sample of over one million French workers from more than five hundred thousand employing firms. We decompose real total annual compensation per worker into components related to observable employee characteristics, personal heterogeneity, firm heterogeneity, and residual variation. Except for the residual, all components may be correlated in an arbitrary fashion. At the level of the individual, we find that person effects, especially those not related to observables like education, are a very important source of wage variation in France. Firm effects, while important, are not as important as person effects. At the level of firms, we find that enterprises that hire high-wage workers are more productive but not more profitable. They are also more capital and high-skilled employee intensive. Enterprises that pay higher wages, controlling for person effects, are more productive and more profitable. They are also more capital intensive but are not more high-skilled labor intensive. We find that person effects explain about 90% of inter-industry wage differentials and about 75% of the firm-size wage effect while firm effects explain relatively little of either differential.

KEYWORDS: Wage determination, person effects, firm effects, inter-industry wage differentials, heterogeneity.

1. INTRODUCTION

FOR DECADES LABOR ECONOMISTS have lamented the lack of microeconomic data relating characteristics of firms to characteristics of their workers (see, for example, Rosen (1986) and Willis (1986)) because such data would permit researchers to begin to disentangle the effects of firm-level decisions from the effects of choices made by workers. Why do high-paying firms provide more than the apparent going wage? Perhaps such a strategy delivers a gain in productivity or profitability that exceeds the incremental wage cost, as predicted by efficiency wage and agency models.[2] Perhaps high-paying firms select workers with higher external wage rates or better firm-specific matches, thus sorting the workers into

[1] The authors gratefully acknowledge the financial and computing support of INSEE. The National Science Foundation supported Abowd and Margolis (SBR 91-11186, 93-21053 and 96-18111). We are grateful for the comments of a co-editor, Ronald Ehrenberg, Hank Farber, Robert Gibbons, Guy Laroque, Stéfan Lollivier, Bentley MacLeod, Olivia Mitchell, Ariel Pakes, Alain Trognon, and Martin Wells as well as for comments received during seminars far too numerous to mention here. The data used in this paper are confidential but the authors' access is not exclusive. Other researchers interested in using these data should contact the Centre de Recherche en Economie et Statistique, INSEE, 15 bd Gabriel Péri, 92245 Malakoff Cedex, France.

[2] See Lazear (1979), Shapiro and Stiglitz (1984), Hart and Hölmstrom (1987), and Sappington (1991) for concise statements of the theories generating these predictions. Tests of these models have been performed by Abowd (1990), Abowd and Kramarz (1993), Cahuc and Dormont (1997), Gibbons and Murphy (1990, 1992), Hutchens (1987), Kahn and Sherer (1990), and Leonard (1990).

firms that have differential observed compensation programs.[3] Although broadly representative linked surveys of firms and workers are not available in the U.S., there have now been numerous studies that attempt to relate firm performance to the design of the compensation system.[4] Furthermore, many have analyzed the inter-industry wage differentials among individuals as if they were the manifestation of differences in firm level compensation policies.[5] In this paper we present the first extensive statistical analysis of simultaneous individual- and firm-level heterogeneity in compensation determination. We examine the variation in personal wage rates holding firm effects constant and variation in firm wage rates holding person effects constant. Due to the matched (person and firm) longitudinal nature of our data, we are able to control for both measured and unmeasured heterogeneity in the workers and their employing firms.

A high-wage worker is a person with total compensation higher than expected on the basis of observable characteristics like labor force experience, education, region, or sex. A high-wage firm is an employer with compensation higher than expected given these same observable characteristics. Until now all empirical analyses of personal and firm heterogeneity in compensation outcomes have relied upon data that were inadequate to identify separately the individual effect necessary to classify a worker as high-wage and the firm effect required to classify a firm as high-wage.

Using a unique longitudinal data set of firms and workers that is representative of private sector French employment, we are able to estimate both person and firm components of compensation determination, allowing for observable and unobservable factors in both dimensions and unrestricted correlation among the effects. Computational complexity prevents full least squares estimation of the models with unobserved heterogeneity in both the person and firm dimensions. After discussing these issues, we examine in detail several related statistical solutions, one of which is a consistent estimator of some of the parameters, and two others that are conditional methods. We also consider other simpler, more classical, techniques in order to assess the importance of person and firm heterogeneity. Although none of these techniques can be used to compute the full least squares solution to the statistical problem, which, for the moment, remains computationally infeasible, all of our methods approximate the full least squares solution and allow the components of person and firm heterogeneity to be intercorrelated. Our consistent method permits estimation of all time-varying coefficients, including those that are heterogeneous. One of our conditional methods, called "order independent," has the advantage that the estimated

[3] This view is espoused by Bulow and Summers (1976), Cain (1976), Jovanovic (1979), and Roy (1951). Weiss and Landau (1984) present a different theoretical version of this model. Some tests include Dickens and Lang (1985), Flinn (1986), Gibbons and Katz (1991), and Heckman and Sedlacek (1985).

[4] See Ehrenberg and Milkovich (1987), Ehrenberg (1990), Ichniowski and Shaw (1993).

[5] See Dickens and Katz (1987), Gibbons and Katz (1992), Groshen (1991), Krueger and Summers (1988), Thaler (1989).

person and firm effects do not depend upon which effect is estimated first and the disadvantage that it cannot impose orthogonality between the estimated residual and the model effects (a characteristic of the full least squares solution). The other conditional method, called "order dependent," has the advantage of imposing this orthogonality but the disadvantage of giving different results depending upon which order is used to estimate the person and firm effects. In particular, the outcome of "persons first and firms second" would differ from "firms first and persons second." In all our estimated models, we find that person effects are statistically more important than firm effects in explaining compensation and performance outcomes and that the two effects are not highly correlated. Using our consistent estimation method, we show empirically that any method in which persons effects are estimated first, whether firm effects are estimated at the same step or after the person effects, performs better than methods in which person effects are estimated after firm effects.

We use our statistical decomposition of wage rates into person and firm effects to address several classic questions in labor economics—the basis for inter-industry wage differentials, the source of the firm size-wage rate relation, the effect of seniority on wage rates, and the relation between pay structure, productivity, and profitability. Surprisingly, our French data give a clear answer to the first question. Virtually all of the inter-industry wage differential is explained by the variation in average individual heterogeneity across sectors. Person effects, and not firm effects, form the basis for most of the inter-industrial salary structure. A very large portion of the positive firm-size wage-rate relation is also due to person effects. The effect of seniority on wage rates is quite heterogeneous across firms; its estimated magnitude is very sensitive to the estimation technique. All our methods for estimating firm effects, including heterogeneous seniority effects, perform well for large firms.

To study pay structure models, we aggregate individual components of compensation to the firm level. Then, we show that firms that hire high-wage workers are more productive per worker, more capital intensive, more professional-employment intensive, and more likely to survive. These same firms are not more profitable, nor are they more skilled-labor intensive. Second, we show that high-wage firms are more profitable, more productive per worker, (possibly) more professional-employment intensive, and (possibly) more capital intensive. High wage firms are unskilled labor intensive and (possibly) less likely to survive.

The paper is organized as follows. In Section 2, we present a detailed motivation of our statistical model in which we relate the different components of our statistical model to wage rate determination models used to study inter-industry wage differentials, firm-size wage effects, the measurement of opportunity wage rates, seniority-wage effects, and the economics of human resource management. In Section 3 we lay out the full details of our statistical methods. We discuss the institutional features of the French labor market and our data sources in Section 4. In Section 5 we discuss our results. Finally, we conclude in Section 6.

2. HETEROGENEITY AND LABOR MARKETS

That labor market outcomes are extremely heterogeneous—observably equivalent individuals earn markedly different compensation and have markedly different employment histories—is one of the enduring features of empirical analyses of labor markets in many countries. This heterogeneity has motivated an enormous literature that attempts to isolate its sources and to identify significant market factors that are statistically related to employment outcomes, particularly earnings or compensation.[6] One strand of this literature has focused on the extent to which wage heterogeneity is related to permanent unmeasured differences among the individuals, what we label a person effect. Another strain of this literature has focused on the extent to which wage heterogeneity is related to permanent differences among the employers, what we label a firm effect.

To put these different models in context, consider the following simple wage equation:

$$y_{it} = \mu_y + (x_{it} - \mu_x)\beta + \theta_i + \psi_{\mathbf{J}(i,t)} + \varepsilon_{it} \tag{2.1}$$

in which y_{it} is the logarithm of annual compensation of individual $i = 1, \ldots, N$ at date $t = 1, \ldots T$; x_{it} is a vector of P time-varying exogenous characteristics of individual i; θ_i is the pure person effect; $\psi_{\mathbf{J}(i,t)}$ is the pure firm effect for the firm at which worker i is employed at date t (denoted by $\mathbf{J}(i,t)$), μ_y is the grand mean of y_{it}, μ_x is the grand mean of x_{it}, and ε_{it} is the statistical residual. Assume that a simple random sample of N individuals is observed for T years.[7] Thus, ε_{it} has the following properties:

$$\mathbf{E}[\varepsilon_{it} | i, t, \mathbf{J}(i,t), x_{it}] = 0$$

and

$$\operatorname{cov}[\varepsilon_{it}, \varepsilon_{ns} | i, t, n, s, \mathbf{J}(i,t), \mathbf{J}(n,s), x_{it}, x_{ns}] = \begin{cases} \sigma_\varepsilon^2 & \text{for } i = n \text{ and } t = s, \\ 0 & \text{otherwise.}^8 \end{cases}$$

[6] See, for example, Rosen (1986), Willis (1986), Becker (1993), Juhn, Murphy, and Pierce (1993), Murphy and Welch (1992), and Blau and Kahn (1996).

[7] The actual data are in the form of an unbalanced panel. For notational simplicity, however, we describe the motivation in terms of a balanced panel. Our complete model is described in the next section and the proofs for the unbalanced case are given in the Statistical Appendix.

[8] One can always allow for a more complicated error structure for ε_{it}; however, as Abowd and Card (1989) show, except for measurement error, this residual exhibits trivial serial correlation in American longitudinal data. Measurement error in the data studied by Abowd and Card, which does exhibit significant serial correlation within individuals, is related to the structure of samples of individuals in which the individuals are the respondents. In this paper, we study data sampled at the level of the individual but reported by the employer; hence, respondent reporting error and other sources of measurement error in individual longitudinal data are not important problems. We will, therefore, maintain the covariance structure assumptions stated here for simplicity. When we consider consistent estimation of β below, we allow for a general covariance structure on ε_{it} for each i.

In matrix notation we have

$$y = X\beta + D\theta + F\psi + \varepsilon, \tag{2.2}$$

where X is the $N^* \times P$ matrix of observable, time-varying characteristics (in deviations from the grand means), D is the $N^* \times N$ matrix of indicators for individual $i = 1, \ldots, N$, F is the $N^* \times mJ$ matrix of indicators for the firm effect at which i works at date t (J firms total),[9] y is the $N^* \times 1$ vector of annual compensation data (also in deviations from the grand mean), ε is the conformable vector of residuals, and $N^* = NT$. The parameters of equation (2.2) are β, the $P \times 1$ vector of coefficients on the time-varying personal characteristics; θ, the $N \times 1$ vector of individual effects;[10] ψ, the $mJ \times 1$ vector of firm effects; and the error variance, σ_ε^2.

Equations (2.1) and (2.2) are interpreted as the conditional expectation of annual compensation given information on the observable characteristics, the date of observation, the identity of the individual, and the identity of the employing firm. The discussion that follows clarifies the interpretation of classical least squares estimates of the parameters β, θ, and ψ when some of these effects are missing or are aggregated into linear combinations. The specification in equation (2.2) is a simplification of the model used in our full analysis below, which we adopt in this section to clarify the discussion. All of the results discussed in this section are general and our Statistical Appendix contains proofs for the general case implemented in our data analyses. As the assumptions on the error process make clear, equations (2.1) and (2.2) impose the assumption of exogenous mobility. In particular, the design matrix for the firm effects, F, is orthogonal to the error process ε. Although endogenous mobility is clearly an important problem, we maintain the assumption of exogenous mobility throughout this paper because we are interested in measuring and summarizing the role of personal and firm heterogeneity in the wage outcomes. The extent to which such heterogeneity arises from endogenous mobility, or other considerations, is the subject of future analyses.

Because many authors have estimated variations of (2.2), but not the full model, there is considerable ambiguity about the interpretation of various combinations of these parameters.[11] Leaving aside the distinction between the

[9] For simplicity in this section we treat the case $m = 1$, so that the firm effect is a constant for each firm. Later in the text we analyze more general firm effects.

[10] The parameter θ includes both the unobservable (to the statistician) individual effect and the coefficients of the non-time-varying personal characteristics.

[11] Since we began working on this paper; several working papers have appeared that use a specification similar to equation (2.2). In particular, see Goux and Maurin (forthcoming), who calculate the exact least squares solution for the modal in equation (2.2) using French data with a much smaller sample of firms and persons than we use; Entorf, Gollac, and Kramarz (forthcoming), who also compute the exact least squares solution using French data with fewer firms and persons than the present paper; and Belzil (1996) and Bingley and Westergård-Nielsen (1996), who use Danish data but do not compute the full least squares solution to equation (2.2); instead, they assume orthogonal firm and person effects. Leonard and Van Audenrode (1996b) use a specification similar to the present one on Belgian data.

conditional and structural interpretation of the parameters, about which we have nothing further to add, it is important to note that the omission or aggregation of one or more of the effects in equation (2.2) can change the meaning of the other effects significantly. Variations in the set of conditioning effects, which give rise to omitted-variable biases, are one source of confusion about the interpretation of the statistical parameters. The use of different linear combinations of the effects in equation (2.2), which gives rise to aggregation biases, is another source of differential interpretations for the parameters. We investigate each of these variations in the parameterization of equation (2.2) in the context of different problems in labor economics.

When the estimated version of equation (2.2) excludes the pure firm effects (ψ), the estimated person effects, θ^*, are the sum of the pure person effects, θ, and the employment-duration weighted average of the firm effects for the firms in which the worker was employed, conditional on the individual time-varying characteristics, X:

$$\theta^* = \theta + (D'M_X D)^{-1} D'M_X F\psi, \tag{2.3}$$

where the notation $M_A \equiv I - A(A'A)^{-1}A'$ for an arbitrary matrix A. Hence, if X were orthogonal to D and F, so that $D'M_X D = D'D$ and $D'M_X F = D'F$, then the difference between θ^* and θ, which is just an omitted variable bias, would be an $N \times 1$ vector consisting, for each individual i, of the employment-duration weighted average of the firm effects ψ_j for $j \in \{\mathbf{J}(i,1), \ldots, \mathbf{J}(i,T)\}$:

$$\theta_i^* - \theta_i = \sum_{t=1}^{T} \frac{\psi_{\mathbf{J}(i,t)}}{T}.$$

The estimated coefficients on the time-varying characteristics in the case of omitted firm effects, β^*, are the sum of the parameters of the full conditional expectation, β, and the omitted variable bias that depends upon the conditional covariance of X and F, given D:

$$\beta^* = \beta + (X'M_D X)^{-1} X'M_D F\psi.$$

Similarly, omitting the pure person effects (θ) from the estimated version of equation (2.2) gives estimates of the firm effects, ψ^{**}, that can be interpreted as the sum of the pure firm effects, ψ, and the employment-duration weighted average of the person effects of all of the firm's employees in the sample, conditional on the time-varying individual characteristics:

$$\psi^{**} = \psi + (F'M_X F)^{-1} F'M_X D\theta. \tag{2.4}$$

Hence, if X were orthogonal to D and F, so that $F'M_X F = F'F$ and $F'M_X D = F'D$, then the difference between ψ^{**} and ψ, again an omitted variable bias, would be a $J \times 1$ vector consisting, for each firm j, of the employment-duration weighted average of the person effects θ_i for $i \in \{\mathbf{J}(i,t) = j \text{ for some } t\}$:

$$\psi_j^{**} - \psi_j = \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\frac{\theta_i 1(\mathbf{J}(i,t) = j)}{N_j} \right]$$

where

$$N_j \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} 1(\mathbf{J}(i,t) = j)$$

and the function $1(A)$ takes the value 1 when A is true and 0 otherwise. The estimated coefficients on the time-varying characteristics in the case of omitted individual effects, β^{**}, are the sum of the parameters of the full conditional expectation, β, and the omitted variable bias that depends upon the covariance of X and D, given F:

$$\beta^{**} = \beta + (X'M_F X)^{-1} X'M_F D\theta. \tag{2.5}$$

Almost all existing analyses of equations like (2.2) produce estimated effects that confound pure person and pure firm effects in a manner similar to those presented above. The possibility of identifying both person and firm effects thus allows us to reexamine many important topics in labor economics using estimates that properly allocate the statistical effects associated with persons and firms.

2.1. *Inter-industry Wage Differentials*

Consider now the analysis of inter-industry wage differentials as done by Dickens and Katz (1987), Krueger and Summers (1987, 1988), Murphy and Topel (1987), Gibbons and Katz (1992), and many others. The principal finding of this literature has been that inter-industrial wage differentials cannot be explained by measured person or firm characteristics. There is continuing controversy regarding the extent to which these differentials are explained by unmeasured person effects, with Krueger and Summers claiming that they are not (Gibbons and Katz concurring), Murphy and Topel claiming that unmeasured person effects are the primary explanation, and Dickens and Katz not able to address the issue. As we make clear in this section, the ability to estimate both person- and firm-level heterogeneity will permit us to substantially resolve this question in our data analysis—in favor of the person-effect explanation, as it turns out.

To standardize notation and parameter interpretation, define the pure inter-industry wage differential, conditional on the same information as in equations (2.1) and (2.2), as κ_k for some industry classification $k = 1, \ldots, K$. Industry is a characteristic of the firm; thus, our definition of the pure industry effect is simply the correct aggregation of the pure firm effects within the industry. We select the definition of an industry effect as the one that corresponds to putting industry indicator variables in equation (2.2) and, then, defining what is left of the pure firm effect as a deviation from the industry effects. Hence, κ_k can be represented as an employment-duration weighted average of the firm effects within the industry classification k:

$$\kappa_k \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\frac{1(\mathbf{K}(\mathbf{J}(i,t)) = k)\psi_{\mathbf{J}(i,t)}}{N_k} \right],$$

where

$$N_k \equiv \sum_{j=1}^{J} 1(\mathbf{K}(j) = k) N_j$$

and the function $\mathbf{K}(j)$ denotes the industry classification of firm j. If we insert this pure industry effect, the appropriate aggregate of the firm effects, into equation (2.1), then the equation becomes

$$y_{it} = x_{it}\beta + \theta_i + \kappa_{\mathbf{K}(\mathbf{J}(i,t))} + (\psi_{\mathbf{J}(i,t)} - \kappa_{\mathbf{K}(\mathbf{J}(i,t))}) + \varepsilon_{it}$$

or, in matrix notation as in equation (2.2),

$$y = X\beta + D\theta + FA\kappa + (F\psi - FA\kappa) + \varepsilon \tag{2.6}$$

where the matrix A, $J \times K$, classifies each of the J firms into one of the K industries; that is, $a_{jk} = 1$ if, and only if, $\mathbf{K}(j) = k$. The parameter vector κ, $K \times 1$, may be interpreted as the following weighted average of the pure firm effects:

$$\kappa \equiv (A'F'FA)^{-1} A'F'F\psi,$$

and the effect $(F\psi - FA\kappa)$ may be re-expressed as $M_{FA}F\psi$. Thus, the aggregation of J firm effects into K industry effects, weighted so as to be representative of individuals, can be accomplished directly by estimation of equation (2.6). Only rank$(F'M_{FA}F)$ firm effects can be separately identified; however, there is neither an omitted variable nor an aggregation bias in the classical least squares estimates of (2.6). To be perfectly clear, equation (2.6) decomposes $F\psi$ into two orthogonal components: the industry effects $FA\kappa$, and what is left of the firm effects after removing the industry effect, $M_{FA}F\psi$.

Authors like Dickens and Katz (1987), Krueger and Summers (1987, 1988), Murphy and Topel (1987), and Gibbons and Katz (1992) do not have information identifying the employing firm, even when they do have longitudinal data.[12] Estimates of industry effects, κ^*, that are computed on the basis of an equation that excludes the remaining firm effects, $M_{FA}F\psi$, are equal to the pure industry effect, κ, plus an omitted variable bias that can be expressed as a function of the conditional variance of the industry effects, FA, given the time-varying charac-

[12] Krueger and Summers (1988, Table V), for example.

teristics, X, and the person effects, D,

$$\kappa^* = \kappa + (A'F'M_{[D\ X]}FA)^{-1}A'F'M_{[D\ X]}M_{FA}F\psi,$$

which simplifies to $\kappa^* = \kappa$ if, and only if, the industry effects, FA, are orthogonal to the subspace $M_{FA}F$, given D and X, which is generally not true even though FA and $M_{FA}F$ are orthogonal by construction.[13] Thus, it is not possible to estimate pure inter-industry wage differentials consistently, conditional on time-varying personal characteristics and unobservable non-time-varying personal characteristics, without identifying information on the underlying firms unless this conditional orthogonality condition holds. Similarly, estimates of the coefficients of the time-varying personal characteristics, β^*, are equal to the true coefficients of the conditional expectation, β, plus an omitted variable bias that depends upon the conditional covariance between these characteristics, X, and the residual subspace of the firm effects, $M_{FA}F$, given D:

$$\beta^* = \beta + (X'M_{[D\ FA]}X)^{-1}X'M_{[D\ FA]}M_{FA}F\psi,$$

which, once again, simplifies to $\beta^* = \beta$ if, and only if, the time-varying personal characteristics, X, are orthogonal to the subspace $M_{FA}F$, given D and FA, which is also not generally true. Thus, it is not possible to estimate the coefficients on time-varying personal characteristics consistently, conditional on industry effects and unobservable non-time-varying personal characteristics, without identifying information on the underlying firms unless this second conditional orthogonality condition holds.

When the estimation of equation (2.6) excludes both person and firm effects, the estimated industry effect, κ_k^{**}, equals the pure industry effect, κ, plus the employment-duration weighted average residual firm effect inside the industry, given X, and the employment-duration weighted average person effect inside the industry, given the time-varying personal characteristics X:

$$\kappa^{**} = \kappa + (A'F'M_X FA)^{-1}A'F'M_X(M_{FA}F\psi + D\theta),$$

which can be restated as

$$\kappa^{**} = (A'F'M_X FA)^{-1}A'F'M_X F\psi + (A'F'M_X FA)^{-1}A'F'M_X D\theta. \tag{2.7}$$

Hence, if industry effects, FA, were orthogonal to time-varying personal characteristics, X, and to non-time-varying personal heterogeneity, D, so that $A'FM_X FA = A'F'FA$, $A'F'M_X F = A'F'F$, and $A'F'M_X D = A'F'D$, the biased inter-industry wage differentials, κ^{**}, would simply equal the pure inter-industry wage differentials, κ, plus the employment-duration-weighted, industry-average

[13] $M_{[D\ X]}$ is the matrix M_Z with $Z \equiv [D \mid X]$ and is not equal to the matrix M_{DX}.

pure person effect, $(A'F'FA)^{-1}A'F'D\theta$, or

$$\kappa_k^{**} = \kappa_k + \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{1[\mathbf{K}(\mathbf{J}(i,t)) = k]\theta_i}{N_k},$$

where $N_k \equiv \Sigma_{i,t} 1[\mathbf{K}(\mathbf{J}(i,t)) = k]$.

Thus, previous analyses that exclude person effects confound the pure inter-industry wage differential with an average of the person effects found in the industry, given the measured personal characteristics, X. To anticipate our results, we use equation (2.7) together with our estimated pure person effects, θ, and our estimated pure firm effects, ψ, to determine what proportion of the estimated inter-industry wage differentials κ^{**} is explained by person effects versus firm effects. We show that the pure inter-industry wage differential, κ, which we interpret, as in this section, as the part due to pure firm effects, is much less important than the contribution of the industry average person effect to κ^{**}.

2.2. *Firm Effects without Personal Heterogeneity*

There is a complementary line of research that attempts to explain heterogeneity in wage rates by using firm effects, for example Groshen (1991, 1996), Davis and Haltiwanger (1996), Entorf and Kramarz (1997, forthcoming) and Kramarz, Lollivier, and Pelé (1996). The principal finding in these studies has been that firm effects are substantially more important than measured personal characteristics in explaining wage variation, even when the measured personal characteristics include detailed occupational effects, which are typically interpreted as a proxy for our pure person effects, θ. An additional conclusion is that the effects of measured personal characteristics, β, are not very sensitive to the inclusion of firm effects. None of the studies in this strain of the wage-determination literature includes both pure person and pure firm effects, as defined in equation (2.1) or (2.2) above.

In our notation, studies like Groshen (1991) estimate ψ^{**}, from equation (2.4), and β^{**}, from equation (2.5). The size of the bias arising from the omission of person effects is, of course, an empirical matter; however, again to anticipate our results, it turns out to be substantial. Most of the estimated firm effect, ψ^{**}, in these studies is due to the employment-duration weighted average of the pure individual effects conditional on X, $(F'M_X F)^{-1}F'M_X D\theta$, and not to the pure firm effect, ψ. Furthermore, the bias in the estimated effects of time-varying personal characteristics, $\beta^{**} - \beta = (X'M_F X)^{-1}X'M_F D\theta$, due to the omission of pure individual effects, is also large.

2.3. *Firm-Size Wage Effects*

The repeated finding of a positive relation between the size of the employing firm and wage rates, even after controlling for a wealth of individual variables (see Brown and Medoff (1989)), has generated many alternative interpretations. Some explanations rely on efficiency wage considerations—monitoring being more difficult in larger firms—or, more generally, upon firm-specific compensation policies.[14] Others rely on the assumed existence of unobserved worker characteristics, compensated by the firms, that only larger firms would be able to spot because of better hiring practices.[15] The estimated firm-size effect on wage rates can be related to what we call pure firm effects as well as to the average person effect within the firm. Using our notation, a firm-size effect, δ, can be modeled using a matrix $S, J \times R$, that maps the size of firm j into R linearly independent functions of its size (polynomials in the logarithm or size intervals, for example). Following the same methods that we used to decompose the inter-industry wage differential, we express the wage equation (2.2) as:

$$(2.8) \qquad y = X\beta + D\theta + FS\delta + M_{FS}F\psi + \varepsilon;$$

so that the pure firm-size effects are related to the underlying pure firm effects by the equation

$$\delta \equiv (S'F'FS)^{-1}S'F'F\psi.$$

Once again, we stress that firm size is a characteristic of the employer; thus, a firm-size effect is simply an aggregation of the pure firm effects and can be analyzed using the same tools that we used for the inter-industry wage differential. Therefore, all of the bias formulas derived for the inter-industry wage differential apply to the problem of estimating the firm-size effects in the presence or absence of the various effects in equation (2.8). In particular, when the firm-size effects are estimated in the presence of measured time-varying personal characteristics, X, and person effects, D, but omitting the remaining firm effects, $M_{FS}F$, the resulting estimated firm-size effects, δ^*, as in Brown and Medoff (1989, Table 2) take the form

$$\delta^* = \delta + (S'F'M_{[D\ X]}FS)^{-1}S'F'M_{[D\ X]}M_{FS}F\psi$$

with a similar equation, which we do not state explicitly, for the bias in the estimation of the parameters β in equation (2.8). The firm-size effects estimated in the absence of firm effects, δ^*, are equal to the pure firm-size effects, δ, if, and only if, firm size, FS, is orthogonal to the residual subspace of firm effects, $M_{FS}F$, given time-varying personal characteristics, X, and person effects, D. As in the case of industry effects, we note that this conditional orthogonality does not follow from the fact that FS and $M_{FS}F$ are orthogonal by construction. Hence, the bias $\delta^* - \delta$ is not generally zero.

Most studies of the firm-size wage effect do not condition on person effects, D. Consequently, the estimated parameter vector associated with the firm-size

[14] See Bulow and Summers (1976), for example.

[15] See Weiss and Landau (1984), for example.

effect in those studies, δ^{**} (in our notation), can be represented as

$$\delta^{**} = (S'F'M_X FS)^{-1} S'F'M_X F\psi + (S'F'M_X FS)^{-1} S'F'M_X D\theta, \tag{2.9}$$

which we interpret as the sum of the firm-size, employment-weighted average firm effect and the similarly-weighted average person effect, conditional on personal characteristics, X. To anticipate our results, again, we use the decomposition displayed in equation (2.9) to explain the relation between firm size and the firm-size class average person and firm effects in our data, conditional on other firm-level and personal variables. The relation between firm size and these components of wage outcomes is, as Brown and Medoff hypothesized, importantly related to both pure firm heterogeneity in compensation, ψ, and pure individual heterogeneity, θ.

2.4. *Measurement of the Internal and External Wage*

Virtually all economic models of labor market outcomes require an estimate of the opportunity cost of the worker's time. In simple, classical equilibrium models without unmeasured person or firm heterogeneity, this generally corresponds to the measured wage rate. In models of wage determination such as quasi-rent splitting[16] or imperfect information (efficiency wage and agency models),[17] unmeasured statistical heterogeneity (person or firm) breaks the direct link between the observed wage rate and the opportunity cost of time. Moreover, such models usually make an explicit distinction between the compensation received and the wage rate available in the employee's next best alternative employment. The statistical model in equation (2.1), while not derived from an explicit labor market model, contains all the observable ele-

[16] In the collective bargaining and wage determination literature, this problem has a long theoretical history (see Leontief (1946), MacDonald and Solow (1981), and most recently Manning (1987)). Many empirical implementations, including Brown and Ashenfelter (1986), Card (1986), MaCurdy and Pencavel (1986), Abowd (1989), Christofides and Oswald (1991, 1992), Nickell and Wadhwani (1991), Abowd and Lemieux (1993), and Blanchflower, Oswald, and Sanfey (1996), use macro-economic wage series or sectoral wage series to represent the opportunity cost of time for the unionized workers. This technique fails to capture important variation in the average personal heterogeneity of the employees of different firms. See Abowd and Kramarz (1993) and Abowd and Allain (1996) for empirical models that permit unobserved heterogeneity in the opportunity cost of time.

[17] For agency models, the theory is summarized in Hart and Holmström (1987) and Sappington (1991). Some empirical implementations include Lazear (1979), Hutchens (1987), Abowd (1990), Gibbons and Murphy (1990, 1992), Jensen and Murphy (1990), Kahn and Sherer (1990), Leonard (1990), Cahuc and Dormont (1997), Cahuc and Kramarz (1997), Kramarz and Rey (1995), and Leonard and Van Audenrode (1996a), all of which require an empirical proxy for the external wage rate in order to identify a component of compensation that is related to performance. See also the summary in Ehrenberg and Milkovich (1987). For efficiency wage models, the theory is summarized in Shapiro and Stiglitz (1984), for the dual labor market version see Bulow and Summers (1976) and Cain (1976). Again, empirical models like Dickens and Lang (1985) require a measure of the opportunity cost in the low-wage sector. The measures used do not allow for unobserved personal heterogeneity between the low and high observed wage groups.

ments from which nonclassical labor market models derive their empirical content. Indeed, the simplest definition of the components of the external and internal wage rate based on a structural model leading to equation (2.1) is given by the following model:

$$(2.10) \qquad y_{it} = x_{it}\zeta + \nu_{it}$$

where $\{x_{it}, \nu_{it}\}$ follows a general stochastic process for $i = 1, \ldots, N$ and $t = 1, \ldots T$ with

$$(2.11) \qquad \mathrm{E}[\{x_{ii}, \nu_{it}\}\{x_{ns}, \nu_{ns}\} \mid i, n, s, t, \mathbf{J}(i,t), \mathbf{J}(n,s)] \neq 0 \quad \text{iff} \quad i = n \quad \text{or} \quad \mathbf{J}(i,t) = \mathbf{J}(n,s).$$

Then,

$$\theta_i = \mathrm{E}[x_{it}\zeta + \nu_{it} \mid i] - \mathrm{E}[x_{it}\zeta + \nu_{it}]$$

and

$$\psi_j = \mathrm{E}[x_{it}\zeta + \nu_{it} \mid \mathbf{J}(i,t) = j] - \mathrm{E}[x_{it}\zeta + \nu_{it}].$$

The model in equation (2.10), together with the assumption (2.11), simply formalizes the conditions under which we can use our maintained assumption of exogenous mobility to apply a structural interpretation to equation (2.1).

2.5. *Analysis of the Seniority-Wage Rate Relation*

In the growing literature on the effects of seniority on wage rates, most authors assume that the relevant coefficient is homogeneous across firms.[18] Ironically, the first uses of the seniority-wage relation to test economic theories (Lazear (1979) and Hutchens (1987)) do not make this assumption. Furthermore, Margolis (1996) has shown, using estimated seniority effects related to those presented in the present paper, that heterogeneity in the returns to seniority is a significant empirical phenomenon and that one's interpretation of the average effect of seniority on wage rates is affected by whether or not the model allows for the heterogeneity. The seniority-wage relation is a firm-specific time-varying effect. Thus, the statistical techniques developed in this paper can be used to model and estimate this effect. We extend the analysis in Margolis (1996) by including a heterogeneous seniority effect in several statistical models. We provide consistent estimates of this effect within firms using assumptions that are comparable to Topel's (1991) assumptions. We compare these results with other estimation techniques that assume heterogeneous or homogeneous seniority effects. Furthermore, we provide direct evidence on the extent to which the between-firm variability in returns to seniority is related to the between-firm variability in initial pay. Several models of lifetime incentive contracts (Becker and Stigler (1974), Lazear (1979)) predict a negative relation, which our statistics support.

[18] See Abraham and Farber (1987), Altonji and Shakotko (1987), Brown (1989), and Topel (1991).

2.6. *Human Resource Management Policies*

In the emerging literature on the economics of human resource management policies (see Ehrenberg (1990) and Lazear (1998)), economists and other organization specialists have argued that a firm's personnel practices, particularly the design of its compensation policy, are directly related to the performance of the firm. These ideas, which we can consider formally in the context of statistical models like equation (2.1), take us back directly to the questions we posed in the introduction. We will measure the opportunity wage of our workers using our estimate of the person-specific heterogeneity in compensation. Thus, at the firm level, the presence of high-wage workers is measured by the average of the person-specific heterogeneity component of pay. The extent to which the firm, through its hiring practices, selects employees who are, on average, better or worse paid than observably-equivalent employees in other firms is, then, directly related to other firm-level outcomes. Again at the firm level, the presence of a high-wage policy will be measured by the firm-specific component of compensation. The extent to which a firm, through its compensation policy, attempts to pay above or below the prevailing market is, then, directly related to other firm-level outcomes. Firm outcomes of interest include the average productivity of labor, sales per employee (as measures of productivity), and the operating income per unit of capital (as a measure of profitability). Existing empirical studies have attempted to relate similar profitability or productivity measures to specific components of the firm's human resource management practices.[19] Because we have a large, representative sample of firms and easily-understood measures of the firms' compensation policies, we are able to supply very direct statistical evidence on the importance of these human resource management practices on the performance and the structure of the firm.

3. STATISTICAL MODEL

3.1. *Specification of the General Model*

Consider, again, our full model as described in equation (2.2). To make our analysis general enough for the data we use, we note that the rows of y, X, D, and F are arranged in the order $i = 1, \dots, N$ and, within each i, $t = n_{i1}, \dots, n_{iT_i}$, where T_i is the total number of years of data available for individual i and the indices $n_{i1}, \dots, n_{iT_i}$ indicate the year corresponding to the first observation on individual i through the last observation on that individual, respectively. Thus

[19] See almost all of the studies in Ehrenberg (1990) but, in particular, Abowd, Hannon, and Milkovich (1990), Kahn and Sherer (1990), and Leonard (1990). Other studies include Weiss and Landau (1984), Ehrenberg and Milkovich (1987), Cahuc and Dormont (1997), Ichniowski and Shaw (1993), Cahuc and Kramarz (1997), Abowd, Kramarz, and Moreau (1996), and Leonard and Van Audenrode (1996b).

the vector y is organized as

$$(3.1) \qquad y = \begin{bmatrix} y_{1,n_{11}} \\ \cdots \\ y_{1,n_{1T_1}} \\ \cdots \\ y_{N,n_{N1}} \\ \cdots \\ y_{N,n_{NT_N}} \end{bmatrix};$$

X, D, F, and ε are organized conformably; and ψ, the parameter vector associated with the firm effects, is $mJ \times 1$ with $m > 1$. To simplify the notation we will refer to a typical element of y as $y_{i,t}$ and a typical element of X, or any similarly organized matrix, as $x_{(i,t),j}$ where the pair (i,t) denotes the row index.

In all of our statistical models, we decompose the person effect, θ_i, into a part that is related to non-time-varying personal characteristics, u_i, and a part that is not observable to the statistician, α_i. We use the orthogonal decomposition of θ_i defined by

$$(3.2) \qquad \theta_i = \alpha_i + u_i \eta$$

where u_i is a vector of non-time-varying measurable personal characteristics, α_i is the person-specific intercept, and η is the vector of coefficients. We also use the following decompositions of ψ_j. The first of these defines a firm effect with $m = 2$,

$$(3.3) \qquad \psi_j = \phi_j + \gamma_j s_{it},$$

where s_{it} denotes individual i's seniority in firm $j = \mathbf{J}(i,t)$ in year t, ϕ_j denotes the firm-specific intercept, and γ_j is the firm-specific seniority coefficient. The second decomposition of ψ_j defines a firm effect with $m = 3$:

$$(3.4) \qquad \psi_j = \phi_j + \gamma_j s_{it} + \gamma_{2j} T_1(s_{it} - 10),$$

where $T_k(x) = 0$ when $x \leq 0$ and $T_k(x) = x^k$ when $x > 0$, and γ_{2j} measures the change in the firm-specific seniority coefficient that occurs after 10 years of seniority. In matrix form equation (3.3) decomposes $F\psi$ as

$$(3.5) \qquad F\psi = F_0 \phi + F_1 \gamma$$

where F_0 is the $N^* \times J$ design matrix associated with the vector of firm specific intercepts, F_1 is the $N^* \times J$ matrix whose columns consist of the direct product of the columns of F_0 and an $N^* \times 1$ vector whose elements are s_{it}, ϕ is the $J \times 1$ vector of firm-specific intercepts and γ is the $J \times 1$ vector of firm-specific seniority coefficients. In matrix notation, equation (3.4) decomposes the firm effect as

$$(3.6) \qquad F\psi = F_0 \phi + F_1 \gamma + F_2 \gamma_2$$

where F_2 is the $N^* \times J$ matrix whose columns consist of the direct product of the columns of F_0 and an $N^* \times 1$ vector whose elements are $\mathbf{T}_1(s_{it} - 10)$, and γ_2

is the $J \times 1$ vector of firm-specific changes in the seniority coefficient after 10 years of seniority.

For completeness, we also note that the derivation of some of our specification tests requires the assumption that $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$. This completes the notation used in the general specification of our statistical model.

3.2. *Identification of Parameters in the General Model*

We now consider basic issues in the identification of the parameters of our model. Although equation (2.2) is just a classical linear regression model, the full design matrix $[X \; D \; F]$ has high column dimension ($N \approx 1{,}000{,}000$ and $J \approx 50{,}000$, estimable). The cross-product matrix

$$\begin{bmatrix} X'X & X'D & X'F \\ D'X & D'D & D'F \\ F'X & F'D & F'F \end{bmatrix}$$

is patterned in the elements $D'D$ and $F'F$; however, projecting onto the columns D leaves a $100{,}000 \times 100{,}000$ unpatterned, nonsparse matrix to invert when $m = 2$ (the linear seniority effect case) because workers move between firms. Indeed, mobility is a necessary condition if one wants to separately identify person effects, θ, and firm effects, ψ, in the general model. Similarly, projecting onto the columns of F leaves a $1{,}000{,}000 \times 1{,}000{,}000$ unpatterned, nonsparse matrix to invert. Clearly, the usual computational methods for least squares estimation of the parameter vector $[\beta' \;\; \theta' \;\; \psi']'$ are not feasible. Hence, because one cannot compute the unconstrained least squares estimates for the model (2.2), we propose several different estimators that attempt to preserve as much of the general structure of the problem as is computationally possible.

Although we do not discuss the origin of our data until Section 4, one aspect of the data, inter-firm mobility, is so critical to the estimation and interpretation of our analyses that we present a summary now. Regardless of the computational approach used, between-employer mobility of the individuals is essential for the identification of our statistical model. Table I examines the pattern of inter-employer movements among all sample individuals. The rows of Table I correspond to the number of years a person is in the sample. The columns, with the exception of column (la), correspond to the number of employers the individual had. An individual contributes to only one cell (again, excepting column (la)). Notice that 59.4% of the individuals in the sample never change employers (column (1)).[20] Approximately one-fifth of the single employer indi-

[20] Notice that the cell (1, 1) contains 318,627 individuals who appear in the sample during a single year. Some of these individuals may represent coding errors in the person identifier; however, it is not possible to correct these errors.

TABLE I

STRUCTURE OF THE INDIVIDUAL DATA BY YEARS IN SAMPLE AND NUMBER OF EMPLOYERS
(Number of Individuals, Most Common Configuration of Employers)

Years in Sample	Number of Employers											Total	Percent
	1	1a	2	3	4	5	6	7	8	9	10		
1	318,627	247,532										318,627	27.3%
	1	1											
2	75,299	57,411	51,066									126,365	10.8%
	2	2	11										
3	46,385	36,540	32,947	19,583								98,915	8.5%
	3	3	12	111									
4	43,019	34,922	26,631	17,191	8,330							95,171	8.2%
	4	4	13	112	1111								
5	41,130	34,596	26,408	15,291	8,685	3,610						95,124	8.2%
	5	5	14	113	1112	11111							
6	29,755	25,388	20,953	13,734	7,592	4,073	1,653					77,760	6.7%
	6	6	15	114	1113	11112	111111						
7	19,413	16,709	17,384	12,039	7,305	3,864	1,931	735				62,671	5.4%
	7	7	16	115	1114	11113	111112	1111111					
8	23,484	20,378	20,421	13,185	7,673	4,001	2,061	917	327			72,069	6.2%
	8	8	44	116	1115	11114	111113	1111112	11111111				
9	38,505	34,147	26,350	15,791	8,590	4,383	2,104	938	362	114		97,137	8.3%
	9	9	54	117	1116	11115	111114	1111113	11111112	111111111			
10	56,881	51,425	32,616	17,728	8,369	3,839	1,837	739	314	109	34	122,466	10.5%
	10[a]	10[a]	64	118	1117	11116	221113	1131112	11111113	111111112	1111111111		
Total	692,498	559,048	254,776	124,542	56,544	23,770	9,586	3,329	1,003	223	34	1,166,305	100.0%
Percent	59.4%	47.9%	21.8%	10.7%	4.8%	2.0%	0.8%	0.3%	0.1%	0.0%	0.0%	100.0%	

Notes: Employment configurations are described in terms of the number of consecutive years spent with each of the individual's employers, in order (e.g. configuration 124 means that the individual spent 1 year with his first employer, then 2 years with his second employer, and finally 4 years with his third employer). Column 1a refers to the subset of individuals with only one employer whose employing firm had at least one other individual who had changed firms at least once in his career (required for least squares identification of both firm and individual effects).
(a) This configuration corresponds to 10 years of data with the first (and only) employer.
Source: Authors' calculations based on the Déclarations Annuelles des Salaires (DAS).

viduals worked in firms with no movers while four-fifths (47.9% of the overall sample, column (1a)) worked in firms that, at one time or another, employed a person who changed employer. Thus, 88.5% of the sample individuals contribute to the estimation of firm-effects. It is also interesting to notice the pattern of employer spells among the movers (columns (2)–(10)). The second line of each cell shows the most frequent configuration of employer spells for individuals in that cell. In almost every case, short spells precede longer spells, indicating that mobility is greater earlier in the career (as Topel and Ward (1992) found for American men). It seems clear from Table I that the data should allow us to separate the individual effect from the firm effect.

3.3. *Identification and Consistent Estimation of β and γ_j*

In this subsection we show how to obtain consistent estimates of β and γ_j using the within-individual-firm differences of the data. This method provides us with our most robust statistical method in the sense that we use no additional statistical assumptions beyond those specified in equation (2.1) and definition (3.3). Consider the first differences:

$$y_{in_{it}} - y_{in_{it-1}} = (x_{in_{it}} - x_{in_{it-1}})\beta + \gamma_{\mathbf{J}(i,n_{it})}(s_{in_{it}} - s_{in_{it-1}}) + \varepsilon_{in_{it}} - \varepsilon_{in_{it-1}} \tag{3.7}$$

for all observations for which $\mathbf{J}(i,n_{it}) = \mathbf{J}(i,n_{it-1})$, which we represent in matrix form as

$$\Delta y = \Delta X\beta + \tilde{F}\gamma + \Delta\varepsilon \tag{3.8}$$

where Δy is $\widetilde{N^*} \times 1$, ΔX is $\widetilde{N^*} \times P$, $\tilde{F}$ is $\widetilde{N^*} \times J$, $\Delta\varepsilon$ is $\widetilde{N^*} \times 1$, and $\widetilde{N^*}$ is equal to the number of (i,t) combinations in the sample that satisfy the condition $\mathbf{J}(i,n_{it}) = \mathbf{J}(i,n_{it-1})$. The matrix $\tilde{F}$ contains the rows of F_1 that correspond to the person-years (i,t) for which the condition $\mathbf{J}(i,n_{it}) = \mathbf{J}(i,n_{it-1})$ is satisfied minus the immediately preceding row. Then,

$$\tilde{\beta} = (\Delta X' M_{\tilde{F}}\, \Delta X)^{-1}\, \Delta X' M_{\tilde{F}}\, \Delta y \tag{3.9}$$

and

$$\tilde{\gamma} = (\tilde{F}'\tilde{F})^{-1}\tilde{F}'\big(\Delta y - \Delta X\tilde{\beta}\big). \tag{3.10}$$

A consistent estimate of $V[\tilde{\beta}]$ is given by

$$\widetilde{\mathrm{V}[\tilde{\beta}]} = (\Delta X' M_{\tilde{F}}\, \Delta X)^{-1}\big(\Delta X' M_{\tilde{F}}\tilde{\Omega} M_{\tilde{F}}\, \Delta X\big)(\Delta X' M_{\tilde{F}}\, \Delta X)^{-1}$$

where

$$\tilde{\Omega} \equiv \begin{bmatrix} \tilde{\Omega}[\Delta\varepsilon_1] & 0 & \cdots & 0 \\ 0 & \tilde{\Omega}[\Delta\varepsilon_2] & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \tilde{\Omega}[\Delta\varepsilon_N] \end{bmatrix}$$

and

$$\tilde{\Omega}[\Delta\varepsilon_i] \equiv \begin{bmatrix} \widetilde{\Delta\varepsilon}^2_{in_2} & \widetilde{\Delta\varepsilon}_{in_2}\widetilde{\Delta\varepsilon}_{in_3} & \cdots & \widetilde{\Delta\varepsilon}_{in_2}\widetilde{\Delta\varepsilon}_{in_{T_i}} \\ \widetilde{\Delta\varepsilon}_{in_3}\widetilde{\Delta\varepsilon}_{in_2} & \widetilde{\Delta\varepsilon}^2_{in_3} & \cdots & \widetilde{\Delta\varepsilon}_{in_3}\widetilde{\Delta\varepsilon}_{in_{T_i}} \\ \cdots & \cdots & \cdots & \cdots \\ \widetilde{\Delta\varepsilon}_{in_{T_i}}\widetilde{\Delta\varepsilon}_{in_2} & \widetilde{\Delta\varepsilon}_{in_{T_i}}\widetilde{\Delta\varepsilon}_{in_3} & \cdots & \widetilde{\Delta\varepsilon}^2_{in_{T_i}} \end{bmatrix}.$$

It is understood that only the rows of $\Delta\varepsilon$ that satisfy the condition $\mathbf{J}(i, n_{it}) = \mathbf{J}(i, n_{it-1})$ are used in the calculation of $\tilde{\Omega}$, which is therefore $\widetilde{N^*} \times \widetilde{N^*}$.[21]

Notice that, given our assumptions, the resulting estimators (3.9) and (3.10) are also unbiased. Our consistent method is not unique but, it has the advantage that the sample on which the estimation is performed includes both workers who remained in the same firm at all dates as well as workers who moved between firms at some point in time during our analysis period. Even for these mobile individuals, all first-differences for which the date t firm differs from the date $t-1$ firm are not included in the estimating sample. Hence, our consistent method is inefficient in the context of the specification of equation (2.2). In addition to this inefficiency, we also note that our consistent method cannot be used to identify separately the firm intercept, ϕ, and the person effect, θ. This results from the restriction of our analysis to a sample based on all observations for which $\mathbf{J}(i, n_{it}) = \mathbf{J}(i, n_{it-1})$. Any method that allows separate identification of the two effects must include in some form the remaining observations. Hence, we turn now to other methods more appropriate to this purpose.

3.4. *Conditional Estimation Methods*

In this section we provide statistical models for estimating all of the effects in equation (2.2) using a class of estimators we call conditional methods because of their relation to standard linear model computational techniques and because of their origins in the panel data literature on person-effect models.[22] Our purpose in developing these methods is to provide estimators that are as similar as possible to the full least squares solution but that are computationally tractable. The basic idea is also simple. Since we cannot compute the full least squares solution, we will have to impose some ancillary orthogonality assumptions in order to proceed. We use information in the data in the form of higher order interactions between observable characteristics, person identity and firm iden-

[21] The formula for the consistent estimator of $V[\tilde{\beta}]$ clearly allows for arbitrary correlation of the residuals ε_{it} over t for each i. Hence, our consistent estimator is unchanged if we permit an arbitrary time-series model for ε_{it}.

[22] The reader familiar with the analysis of variance as considered in, for example, Scheffé (1959) and Searle et al. (1992), will notice that our conditional methods can also be derived as analysis of covariance models in which the data are adjusted to remove certain effects, our conditioning variables, before the conventional analysis of covariance formulas are applied to the model.

tity, which are excluded by hypothesis from equation (2.2), to proxy for the correlation between X, D, and F. Then, we impose conditional orthogonality, given these higher order interactions. Since we have a consistent, but inefficient, estimator of some of the effects, we will use that estimator to assess the quality of our conditional estimation methods when we consider formal specification checks. We will, thus, have some formal and some informal methods for comparing a variety of estimators, none of which is the full least squares solution for estimating the parameters of equation (2.2).

Consider a matrix of variables $Z, N^* \times Q$, which depends upon Q functions of the information in X, D, and F. Using conditional methods we calculate the least squares estimates of equation (2.2) under different maintained hypotheses about the conditional orthogonality of X, D, and F, given Z. The first of these hypotheses imposes that the effects X and D be orthogonal to the projection of F onto the null space of Z. Under this hypothesis the basic equation can be restated as

$$y = X\beta + D\theta + Z\lambda + M_Z F\psi + \varepsilon \tag{3.11}$$

where the auxiliary parameter $\lambda \equiv (Z'Z)^{-1}Z'F\psi$. The assumption of conditional orthogonality between X and F, given Z, and between D and F, given Z, implies that

$$X'M_Z F = 0 \tag{3.12}$$

and

$$D'M_Z F = 0. \tag{3.13}$$

Hence, the conventional least squares formula for the estimator of the original parameters, $[\,\beta' \;\; \theta' \;\; \psi'\,]'$, and the auxiliary parameters, λ, is

$$\begin{bmatrix} \hat{\beta} \\ \hat{\theta} \\ \hat{\lambda} \\ \hat{\psi} \end{bmatrix} = \begin{bmatrix} X'X & X'D & X'Z & X'M_Z F \\ D'X & D'D & D'Z & D'M_Z F \\ Z'X & Z'D & Z'Z & Z'M_Z F \\ F'M_Z X & F'M_Z D & F'M_Z Z & F'M_Z F \end{bmatrix}^{-} \begin{bmatrix} X'y \\ D'y \\ Z'y \\ F'M_Z y \end{bmatrix} \tag{3.14}$$

where the notation $[\;]^-$ denotes a generalized inverse.[23] Since the elements $X'M_Z F$, $D'M_Z F$, and $Z'M_Z F$ are zero, either by hypotheses (3.12) and (3.13) or by construction, the formula (3.14) can be restated as

$$\begin{bmatrix} \hat{\beta} \\ \hat{\theta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} X'X & X'D & X'Z \\ D'X & D'D & D'Z \\ Z'X & Z'D & Z'Z \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ D'y \\ Z'y \end{bmatrix} \tag{3.15}$$

and

$$\hat{\psi} = (F'M_Z F)^{-} F'M_Z y. \tag{3.16}$$

[23] The use of a g-inverse is required because $(F'M_Z F)$ is rank $mJ - 1 - Q$.

As we demonstrated in Section 3.3, certain parameters in our model can be estimated consistently without the use of ancillary assumptions like equations (3.12) and (3.13). Consistent estimation of other parameters requires some extra hypotheses. If the conditional methods work well, then the conditional estimates of β and γ_j should not be too far from the estimates produced by the consistent method. This insight is the basis for the specification checks that we derive below.

3.4.1. *Order-independent estimation*

Our first method for the computation of the solution to equations (3.15) and (3.16) can be accomplished in two steps, which can be performed in either order, hence our designation of this method as "order independent." In the first step, called the within-D step, the parameters in equation (3.15) are estimated by conventional longitudinal methods in which X and Z are projected on D to produce the estimates of β and λ given by

$$(3.17) \qquad \begin{bmatrix} \hat{\beta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} X'M_D X & X'M_D Z \\ Z'M_D X & Z'M_D Z \end{bmatrix}^{-1} \begin{bmatrix} X'M_D y \\ Z'M_D y \end{bmatrix},$$

which are usually called the "within-person" estimators of these parameters. The associated estimator of θ is

$$(3.18) \qquad \hat{\theta} = (D'D)^{-1} D'\left(y - X\hat{\beta} - Z\hat{\lambda}\right).$$

The second step in the computation of the complete set of order-independent, conditional least squares estimates for equation (3.11), called the within-F step, requires the solution of equation (3.16). This is accomplished by computing the least squares estimates of the parameters in the regression of y on F and Z jointly:

$$(3.19) \qquad y = F\psi + Z\pi + \upsilon,$$

where ψ is the same parameter vector that appears in equation (2.2), π is a $Q \times 1$ vector of auxiliary parameters, and $\upsilon \sim N(0, \sigma_\upsilon^2 I)$ because of the conditional orthogonality conditions imposed in equations (3.12) and (3.13). Computation of $\hat{\psi}$ is accomplished in two steps that are directly analogous to the method used in equations (3.17) and (3.18). First, compute $\hat{\pi}$ by the within-F estimator

$$(3.20) \qquad \hat{\pi} = (Z'M_F Z)^{-1} Z'M_F y.$$

Then, compute $\hat{\psi}$ with the estimator

$$(3.21) \qquad \hat{\psi} = (F'F)^{-1} F(y - Z\hat{\pi}).$$

The proof that the formulas (3.16) and (3.21) are equivalent follows. First, note that

$$\hat{\pi} = \left[(Z'Z)^{-1} + (Z'Z)^{-1}Z'F(F'M_ZF)^{-}F'Z(Z'Z)^{-1}\right]Z'y$$
$$-(Z'Z)^{-1}Z'F(F'M_ZF)^{-}F'y$$

by direct application of the partitioned inverse formula to the full least squares solution to equation (3.19). Hence,

$$(3.22) \qquad y - Z\hat{\pi} = \left[M_Z - P_ZF(F'M_ZF)^{-}F'P_Z + P_ZF(F'M_ZF)^{-}F'\right]y$$

where $P_Z \equiv I - M_Z$. Substituting equation (3.22) into equation (3.21) yields

$$\begin{aligned}\hat{\psi} &= (F'F)^{-1}F'(y - Z\hat{\pi})\\ &= (F'F)^{-1}F'M_Zy - (F'F)^{-1}F'P_ZF(F'M_ZF)^{-}F'P_Zy\\ &\quad + (F'F)^{-1}F'P_ZF(F'M_ZF)^{-}F'y\\ &= \left[(F'F)^{-1}(F'M_ZF) + (F'F)^{-1}F'P_ZF\right](F'M_ZF)^{-}F'M_Zy\\ &= (F'M_ZF)^{-}F'M_Zy. \qquad\qquad Q.E.D.\end{aligned}$$

In some applications, the matrix F is just F_0, the design matrix for a single firm-specific effect ($m = 1$), and the computation of equations (3.20) and (3.21) can be accomplished by conventional formulas in which the values of y and Z are deviated from within-firm means in order to compute $\hat{\pi}$. In our estimation using this conditional model we let $m = 2$ in order to capture a firm-specific intercept and seniority slope according to equation (3.5). The within-F step regression becomes

$$(3.23) \qquad y = F_0\phi + F_1\gamma + Z\pi + \nu.$$

In estimation of the firm effects by equation (3.23), the computation of $\hat{\pi}$, $\hat{\phi}$, and $\hat{\gamma}$ is more complex than for the case in which $F = F_0$. These complexities are described in the Statistical Appendix.

The estimation of the correct covariance matrix for the combined within-D and within-F estimator requires calculation of the correct residual for the full model in (2.2),

$$\hat{\varepsilon} = \left(y - X\hat{\beta} - D\hat{\theta} - Z\hat{\lambda} - M_ZF\hat{\psi}\right).$$

The computation of this residual is not straightforward. The first part of the residual is computed at the within-D step as

$$\hat{\varepsilon}^{[1]} = \left(y - X\hat{\beta} - D\hat{\theta} - Z\hat{\lambda}\right).$$

The second part of the residual is computed at the end of the within-F step as

$$\hat{\varepsilon}^{[2]} = M_ZF\hat{\psi} = F\hat{\psi} - Z\tilde{\lambda}$$

where $\tilde{\lambda} \equiv (Z'Z)^{-1}Z'F\hat{\psi}$. Finally,

$$\hat{\varepsilon} = \hat{\varepsilon}^{[1]} - \hat{\varepsilon}^{[2]}.$$

The standard analysis of variance estimator for the variance of the residual ε is given by

$$(3.24) \qquad \hat{\sigma}_\varepsilon^2 = \frac{\left(y - X\hat{\beta} - D\hat{\theta} - Z\hat{\lambda} - M_Z F\hat{\psi}\right)'\left(y - X\hat{\beta} - D\hat{\theta} - Z\hat{\lambda} - M_Z F\hat{\psi}\right)}{N^* - P - N - Q - (mJ - 1 - Q)}$$

where we note explicitly that the estimation of the mJ firm effects uses only $mJ - Q - 1$ degrees of freedom and that the Q degrees of freedom missing from the firm effects have been used to estimate $\hat{\lambda}$. The proof follows:

$$\hat{\varepsilon} = \left[I - W(W'W)^{-1}W' - M_Z F(F'M_Z F)^{-1}F'M_Z\right]\varepsilon = M_{[W\ M_Z F]}\varepsilon$$

where $W \equiv [X\ D\ Z]$. Under the maintained orthogonality conditions in equations (3.12) and (3.13), the quadratic form

$$\frac{\varepsilon'\varepsilon}{\sigma_\varepsilon^2} \sim \chi^2_{N^*}.$$

and

$$(3.25) \qquad \frac{\varepsilon'\varepsilon}{\sigma_\varepsilon^2} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_\varepsilon^2} + \frac{\varepsilon' W(W'W)^{-1}W'\varepsilon}{\sigma_\varepsilon^2} + \frac{\varepsilon' M_Z F(F'M_Z F)^{-1}F'M_Z\varepsilon}{\sigma_\varepsilon^2}$$

Since $W(W'W)^{-1}W'M_Z F(F'M_Z F)^{-1}F'M_Z = 0$, the last two quadratic forms on the right-hand side of equation (3.25) are independent χ^2 random variables with $\text{rank}[W(W'W)^{-1}W'] = P + N + Q$ and $\text{rank}[M_Z F(F'M_Z F)^{-1}F'M_Z] = mJ - Q - 1$ degrees of freedom, respectively. Thus,

$$\frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_\varepsilon^2} \sim \chi^2_{N^* - P - N - mJ + 1}.$$

The error degrees of freedom for the complete model is, thus, $N^* - P - N - mJ + 1$, so that the dimensionality of the auxiliary parameter vector λ does not affect the goodness of fit of the model in equation (3.11).

3.4.2. *Order-dependent estimation*

The order-dependent method, conditional on Z, means that the estimation of certain effects is performed before others; that is, that the residuals from the first step are used to compute the estimates of the second step. The result is order-dependent because estimating person effects before firm effects is not the same as estimating firm effects before person effects. We describe in detail the order-dependent: persons first method. We comment only briefly on the order-dependent: firms first method, because the analogous formulas are straightforward.

The first step of our order-dependent: persons first method uses the same conditional estimation equations that were described above for the order-independent method to estimate the coefficients of the time-varying observable variables, β, the person effects, θ, and the conditioning effects, λ, the coefficient of the variables Z. This is done according to equation (3.15). Hence, the estimated coefficients are given by equations (3.17) and (3.18).

In the second step of the order-dependent: persons first method, we estimate the firm effects using equation (3.4) and its matrix specification (3.6).[24] Define

$$(3.26) \qquad \{j\} \equiv \{(i,t) \mid \mathbf{J}(i,t) = j\}, \quad \text{a set with } N_j \text{ elements.}$$

Now,

$$(3.27) \qquad \hat{y}_{\{j\}} \equiv y_{\{j\}} - x_{\{j\}}\hat{\beta} - \hat{\theta}_{\{j\}},$$

where

$$(3.28) \qquad y_{\{j\}} \equiv \begin{bmatrix} \cdots \\ y_{ns} \\ \cdots \end{bmatrix}, \quad \forall (n,s) \in \{j\},$$

and similarly for $x_{\{j\}}$ and $\hat{\theta}_{\{j\}}$. Equations (3.26) and (3.27) group all of the observations on individuals employed by the same firm into the vector $\hat{y}_{\{j\}}$, which is expressed as the deviation from the first-step estimated $x\hat{\beta}$ and $\hat{\theta}$. The firm-level equation is

$$(3.29) \qquad \hat{y}_{\{j\}} = F_{\{j\}} \begin{bmatrix} \phi_j \\ \gamma_j \\ \gamma_{2j} \end{bmatrix} + \zeta_{\{j\}}$$

where

$$(3.30) \qquad F_{\{j\}} \equiv \begin{bmatrix} & \cdots & \\ 1 & s_{ns} & \mathrm{T}_1(s_{ns} - 10) \\ & \cdots & \end{bmatrix}, \quad \forall (n,s) \in \{j\}$$

and

$$(3.31) \qquad \zeta_{\{j\}} \equiv \varepsilon_{\{j\}} + x_{\{j\}}\left(\beta - \hat{\beta}\right) + \left(\theta_{\{j\}} - \hat{\theta}_{\{j\}}\right).$$

[24] Since this second method is much simpler to implement than the first one, we use a specification of the firm effect that is more complicated by including a linear spline after 10 years of seniority.

Least squares estimation of (3.29) yields the estimator

$$\begin{bmatrix} \hat{\phi}_j \\ \hat{\gamma}_j \\ \hat{\gamma}_{2j} \end{bmatrix} = (F'_{\{j\}}F_{\{j\}})^{-1}F'_{\{j\}}\hat{y}_{\{j\}}, \quad \text{for } j = 1, \ldots J. \tag{3.32}$$

The asymptotic distribution of the estimator in equation (3.32) is

$$\begin{bmatrix} \hat{\phi}_j \\ \hat{\gamma}_j \\ \hat{\gamma}_{2j} \end{bmatrix} \rightarrow \mathrm{N}\left(\begin{bmatrix} \phi_j \\ \gamma_j \\ \gamma_{2j} \end{bmatrix}, \Omega_j \right) \quad \text{as} \quad N_j \rightarrow \infty \tag{3.33}$$

where

$$\Omega_j \equiv \sigma_\varepsilon^2 (F'_{\{j\}}F_{\{j\}})^{-1} + (F'_{\{j\}}F_{\{j\}})^{-1}F'_{\{j\}} \left(X_{\{j\}}\mathrm{var}\left[\hat{\beta}\right]X'_{\{j\}} + \mathrm{var}\left[\hat{\theta}_{\{j\}}\right] + 2X_{\{j\}}\mathrm{cov}\left[\hat{\beta}, \hat{\theta}_{\{j\}}\right]\right) F_{\{j\}}(F'_{\{j\}}F_{\{j\}})^{-1}. \tag{3.34}$$

The first step of our order-dependent "firms first" method begins with the equations (3.16) or (3.21) defining the estimator $\hat{\psi}$ in the order-independent method. The order-dependent "firms first" estimator for β and θ is based on conventional computational formulas applied to equation

$$y - F\hat{\psi} = X\beta + D\theta + \xi$$

where $\xi = \varepsilon + F(\psi - \hat{\psi})$. The order-dependent "firms first" estimators are

$$\hat{\beta} = (X'M_D X)^{-1} X'M_D\left(y - F\hat{\psi}\right) \tag{3.35}$$

and

$$\hat{\theta} = (D'D)^{-1}D'\left(y - X\hat{\beta} - F\hat{\psi}\right). \tag{3.36}$$

The asymptotic covariance matrices of the estimators in equations (3.35) and (3.36) can be derived directly from the standard formulas and the asymptotic covariance matrix of the order-dependent "firms first" estimator of ψ, which is just $\sigma_\varepsilon^2(F'M_Z F)^-$.

3.4.3. *Relation between the order-independent and order-dependent estimates*

In the discussion of our empirical results we refer repeatedly to different forms of the conditional estimators. In this subsection we summarize the relations among the different conditional estimators. The order-independent estimator for β and θ, equations (3.17) and (3.18), is identical to the order-dependent "persons first" estimator for β and θ. The order-dependent "persons first" estimator for ψ is equation (3.32). The order-independent estimator for ψ, equation (3.16) or (3.21), is identical to the order-dependent "firms first" estimator for ψ. The order-dependent "firms first" estimators for β and θ are equations (3.35) and (3.36).

3.4.4. *Estimation of components of the individual effect*

Regardless of the estimator used for θ_i, we also decompose the individual effect into a component attributable to fixed individual characteristics, u_i (such as education), and an unobservable component, α_i, as shown in equation (3.2). To recover the α_i and $u_i\eta$ parts of the individual effect, we use the estimated individual effects, $\hat{\theta}_i$, and their associated estimated sampling variances to estimate the equation (3.2) by generalized least squares. We obtain $\hat{\eta}$, which satisfies

$$(3.37)\qquad \hat{\eta} \to \mathrm{N}\left(\eta, \left(U'\mathrm{diag}\left(\mathrm{var}\left[\hat{\theta}_i\right]\right)^{-1} U\right)^{-1}\right) \quad \text{as} \quad N \to \infty$$

where

$$(3.38)\qquad U \equiv \begin{bmatrix} u_1 \\ \cdots \\ u_N \end{bmatrix}$$

and diag(var[$\hat{\theta}_i$]) is a diagonal matrix $[\hat{\sigma}_\varepsilon^2/T_i]$, the asymptotic variances of $\hat{\theta}_i$ using the residual variance estimator from equation (3.24). The estimator of α_i is

$$(3.39)\qquad \hat{\alpha}_i = \hat{\theta}_i - u_i\hat{\eta}$$

and is unbiased and asymptotic in T_i (Chamberlain (1984)). We show below that statistics based upon aggregating $\hat{\theta}_i$ and $\hat{\alpha}_i$ to the level of the firm are consistent.

3.5. *Specification Checks*

Because of the result in equation (3.25), namely that the goodness of fit of the model does not depend upon the number of auxiliary parameters used in the within-D or within-F step, conventional specification tests and Bayesian model selection procedures are not applicable. Essentially, we must maintain the conditional orthogonality assumptions (3.12) and (3.13) in order to compute any estimates at all of equation (3.11). Although we cannot compute a classical specification test in the sense of Hausman (1978), we can use those principles to derive a specification test whose distribution is known under the null hypothesis

$$H_0 : \lambda = (Z'Z)^{-1} Z'F\psi,$$

which is the definition of the auxiliary parameter λ under the conditional orthogonality hypotheses of equations (3.12) and (3.13).

Consider the residual from equation (2.2) when $\hat{\beta}$, $\hat{\theta}$, and $\hat{\psi}$ are defined according to the order-independent estimation formulas (3.17), (3.18), and (3.16), respectively,

$$\begin{aligned}\tilde{\varepsilon} &= \left(y - X\hat{\beta} - D\hat{\theta} - F\hat{\psi}\right)\\ &= \varepsilon - X\left(\hat{\beta} - \beta\right) - D(\hat{\theta} - \theta) - Z(\hat{\lambda} - \lambda) - M_Z F\left(\hat{\psi} - \psi\right)\\ &\quad + Z(\hat{\lambda} - \lambda) - P_Z F\left(\hat{\psi} - \psi\right)\\ &= \hat{\varepsilon} + Z(\hat{\lambda} - \lambda) - P_Z F\left(\hat{\psi} - \psi\right)\\ &= \hat{\varepsilon} + Z\left(\hat{\lambda} - (Z'Z)^{-1} Z'F\hat{\psi}\right) - Z\left(\lambda - (Z'Z)^{-1} Z'F\psi\right).\end{aligned}$$

Hence, under the null hypothesis

$$\tilde{\varepsilon} - \hat{\varepsilon} = Z(\hat{\lambda} - \tilde{\lambda}).$$

The statistic $\hat{\lambda} - \tilde{\lambda}$ is very similar to a specification test statistic since it is the difference between the Cramer-Rao efficient estimator of λ, namely $\hat{\lambda}$, and an inefficient but unbiased estimator of the same auxiliary parameter, namely $\tilde{\lambda}$. By direct application of the Cramer-Rao lower bound implied by the efficiency of $\hat{\lambda}$ for the model given in equation (3.11), we have

$$\hat{\lambda} - \tilde{\lambda} \sim N(0, \sigma_\varepsilon^2 \Omega)$$

where

$$\sigma_\varepsilon^2 \Omega \equiv (Z'Z)^{-1} Z'F \operatorname{var}\left[\hat{\psi}\right] F'Z(Z'Z)^{-1} - \operatorname{var}[\hat{\lambda}],$$

and $\operatorname{var}[\hat{\lambda}]$ and $\operatorname{var}[\hat{\psi}]$ are the covariance matrices of the parameter estimators $\hat{\lambda}$ and $\hat{\psi}$, respectively, as computed in the solutions to equations (3.15) and (3.16). The variance of $\hat{\lambda} - \tilde{\lambda}$ is guaranteed to be positive semi-definite by the efficiency of $\hat{\lambda}$. Thus, a test of the specification of the model can be based upon the distribution of $\tilde{\varepsilon} - \hat{\varepsilon}$. The statistic

$$\frac{(\tilde{\varepsilon} - \hat{\varepsilon})'[Z\Omega Z']^{-}(\tilde{\varepsilon} - \hat{\varepsilon})}{\sigma_\varepsilon^2} \sim \chi^2_{Q^*},$$

where $Q^* = \operatorname{rank}[Z\Omega Z'] \leq Q$. An equivalent statistic that is easier to compute is based on the distribution of $Z'(\tilde{\varepsilon} - \hat{\varepsilon})$, a $Q \times 1$ random vector:

$$\begin{aligned}&\frac{(\tilde{\varepsilon} - \hat{\varepsilon})' Z(Z'Z)^{-1} \Omega^{-} (Z'Z)^{-1} Z'(\tilde{\varepsilon} - \hat{\varepsilon})}{\sigma_\varepsilon^2}\\ &\quad = \frac{(\hat{\lambda} - \tilde{\lambda})'(Z'Z)^{-1} \Omega^{-} (Z'Z)^{-1}(\hat{\lambda} - \tilde{\lambda})}{\sigma_\varepsilon^2} \sim \chi^2_{Q^*},\end{aligned} \tag{3.40}$$

where $Q^* = \operatorname{rank}[\Omega] \leq Q$. This statistic is the only formal specification test that we derive for the reasonableness of the set of conditioning variables, Z.

To compare the different estimators of the β coefficients, we rely on our consistent estimate of β and use Hausman (1978) statistics. We derive conventional specification tests of the difference between the consistent estimates of β based on equation (3.9) and estimates from other methodologies, including our conditional methods.

3.6. *The Construction of Z*

The role of the conditioning variables, Z, is to proxy for the covariation among the effects represented by X, D, and F. The columns of Z should be chosen to preserve as many of the effects ψ as possible, recalling that each column of Z reduces the rank of $(F'M_Z F)$ by one, while capturing as much of the conditional covariance of X and D with F as possible. Since these are competing goals, we will rely on judgement and on the specification test in equation (3.40) to choose a reasonable set of Z variables. We begin by noting that every column of Z increases the computational complexity of solving the equation system (3.15) and (3.16) in proportion to N^*Q^2 in terms of both storage and calculations. It is therefore necessary to accept some a priori restrictions on this auxiliary design matrix Z. Second, we note that the within-F regression in our conditional estimation procedure will not be well-defined for Z variables that do not have within-F variation. In order to give all columns of Z some within-F variation while, at the same time, inducing correlation with X and D, we chose the Z variables as interactions between firm characteristics (functions of F) and personal characteristics (functions of X and D). Under the specification described by equation (3.11), none of these interactions enters the model directly.

The columns of Z are defined as follows. Let

$$\bar{x}_i \equiv \frac{\sum_{t=1}^{T_i} x_{in_{it}}}{T_i} = \text{the within-person mean of } x_{it},$$

and

$$\bar{f}_j \equiv \frac{\sum_{(i,t) \in \{\mathbf{J}(i,t)=j\}} f_{(i,t)j}}{N_j} = \text{firm average of characteristic } f_{(i,t)j},$$

where the firm characteristics are measured by taking functions of the columns of F. In particular, firm size can be measured as a fixed constant times the number of person-years observed in firm j over the life of the sample, N_j.[25] The industry of firm j can be determined by applying a classification matrix A, $J \times K$

[25] We can calculate firm size in our sample using the following method. In our data, the employee sampling rate is $1/25$th and the number of at risk years is 10; hence, the constant $= 2.5$. Thus, in matrix form, we convert F_0 into a vector of firm sizes, L, as

$$L = F_0[e'_{N*} F_0 \cdot 2.5]'$$

where e_{N*} is an $N^* \times 1$ vector of ones.

to F_0 so that the result $F_0 A$ classifies each row of F_0, thus all N^* persons-years, into one of K unique industries. The firm characteristics actually used in our analysis are firm size, its square, and a 10-industry classification. The personal characteristics actually used were labor force experience (time-varying) and age at the end of schooling (non-time-varying). The rows of Z were constructed as

$$\text{row}\,(i,t) = \begin{bmatrix} \bar{x}_i & u_i \end{bmatrix} \otimes \begin{bmatrix} \bar{f}_{\mathbf{J}(i,t)} \end{bmatrix} \otimes \begin{bmatrix} 1 & s_{it} \end{bmatrix}.$$

3.7. *Analysis of Firm-level Outcomes*

Our analysis of firm-level outcomes requires summary statistics, by firm, of the effects estimated from equation (2.2). Although we use several different estimators for these effects, we always use the same aggregation formulas; so, we have shown those formulas using generic estimators for the underlying parameters.

First consider firm-level averages of the person effects θ_i and α_i,

$$\text{(3.41)} \qquad \hat{\theta}_j \equiv \frac{1}{N_j} \sum_{(i,t)\in\{\mathbf{J}(i,t)=j\}} \hat{\theta}_i$$

and

$$\hat{\alpha}_j \equiv \frac{1}{N_j} \sum_{(i,t)\in\{\mathbf{J}(i,t)=j\}} \hat{\alpha}_i .$$

We use the asymptotic distribution for $\hat{\alpha}_j$:

$$\text{(3.42)} \qquad \hat{\alpha}_j \to \mathrm{N}\left(\alpha_j, \sigma^2_{\alpha_j}\right), \quad \text{as} \quad N_j \to \infty$$

where

$$\sigma^2_{\alpha_j} \equiv \frac{1}{N_j^2} \sum_{i=1}^{N_j} \frac{\sigma_\varepsilon^2}{T_i} \left[1 - \frac{T_i}{\sigma_\varepsilon^2} u_i' \left(U' \operatorname{diag}\left(\operatorname{var}\left[\hat{\theta}_i\right]\right)^{-1} U \right)^{-1} u_i \right]$$

and for $\hat{\theta}_j$ (not shown).[26] Similarly, the firm-level average education effect is given by

$$\text{(3.43)} \qquad \bar{u}_j \hat{\eta} \equiv \frac{1}{N_j} \sum_{(i,t)\in\{\mathbf{J}(i,t)=j} u_i \hat{\eta}$$

with asymptotic distribution based upon (3.37). In all our asymptotic results we hold constant the distribution of firm sizes. Thus as $N, N_j \to \infty$, we assume that their ratio N_j/N goes to a nonzero constant.

[26] The formula for the symptotic distribution of $\hat{\theta}_j$ is identical to the one for $\hat{\alpha}_j$ with the quadratic form in u_i removed.

We consider next the statistical relation between firm-level outcomes and our measures of firm-level compensation policy. Our basic model is

$$(3.44)\qquad p_j = \begin{bmatrix} \alpha_j & \bar{u}_j\eta & \phi_j & \gamma_j & \gamma_{2j} & q_j \end{bmatrix}\begin{bmatrix} \zeta \\ \rho \end{bmatrix} + \xi_j$$

where $j = 1, \ldots J$, p_j is any firm-level outcome, $[\alpha_j \;\; \bar{u}_j\eta \;\; \phi_j \;\; \gamma_j \;\; \gamma_{2j}]$ is a vector of firm-level compensation measures, ζ is a vector of parameters of interest, q_j is a vector of other firm-level variables, ρ is a vector of parameters associated with q_j, and ξ_j is a zero-mean homoscedastic statistical error.[27] In the regression analysis, firm-level outcomes and firm-level compensation variables were measured using data from two independently drawn samples. However, the firm-level compensation variables derived from our individual sample are estimated regressors. Consequently, we must allow for the estimation errors in $\hat{\alpha}_j$, $\bar{u}_j\hat{\eta}$, $\hat{\phi}_j$, $\hat{\gamma}_j$, and $\hat{\gamma}_{2j}$ in our assessment of the precision of the estimation of firm-level equations.[28] Equation (3.44) becomes

$$(3.45)\qquad \begin{aligned} p_j = {} & \begin{bmatrix} \hat{\alpha}_j & \bar{u}_j\hat{\eta} & \hat{\phi}_j & \hat{\gamma}_j & \hat{\gamma}_{2j} & q_j \end{bmatrix}\begin{bmatrix} \zeta \\ \rho \end{bmatrix} \\ & + \Big(\begin{bmatrix} \alpha_j & \bar{u}_j\eta & \phi_j & \gamma_j & \gamma_{2j} \end{bmatrix} \\ & - \begin{bmatrix} \hat{\alpha}_j & \bar{u}_j\hat{\eta} & \hat{\phi}_j & \hat{\gamma}_j & \hat{\gamma}_{2j} \end{bmatrix} \Big)\zeta + \xi_j \end{aligned}$$

where $([\alpha_j \;\; \bar{u}_j\eta \;\; \phi_j \;\; \gamma_j \;\; \gamma_{2j}] - [\hat{\alpha}_j \;\; \bar{u}_j\hat{\eta} \;\; \hat{\phi}_j \;\; \hat{\gamma}_j \;\; \hat{\gamma}_{2j}])\zeta$ is the error associated with the first-step estimation of the firm-level compensation measures.[29] In order to derive the error covariance matrix for equation (3.45), let

$$P_j'\big(\hat{\delta}_j\big) \equiv \begin{bmatrix} \hat{\alpha}_j & \bar{u}_j\hat{\eta} & \hat{\phi}_j & \hat{\gamma}_j & \hat{\gamma}_{2j} & q_j \end{bmatrix}$$

and

$$\hat{\delta}_j' \equiv \begin{bmatrix} \hat{\alpha}_j & \bar{u}_j\hat{\eta} & \hat{\phi}_j & \hat{\gamma}_j & \hat{\gamma}_{2j} \end{bmatrix}.$$

[27] This is the most general specification, corresponding to the parameterization of the firm effect ($m = 3$) used in our order-dependent "persons first" method. In some of our firm-level analyses the terms involving γ_{2j} do not appear because the underlying firm effects were of lower dimension ($m = 2$).

[28] The firm-level regressor $\bar{x}_j\hat{\beta}$ also contains some measurement error, in principle; however, the vector $\hat{\beta}$ is estimated with such precision that we do not carry along its estimated covariance matrix (including its estimated covariance matrix with $\hat{\alpha}_j$, $\bar{u}_j\hat{\eta}$, $\hat{\phi}_j$, $\hat{\gamma}_j$, and $\hat{\gamma}_{2j}$) in these calculations. Hence, we place $\bar{x}_j\hat{\beta}$ in the list of q_j.

[29] We adopt the model of Pagan (1984); namely, that the regression of interest relates a function of the individual-level data and several firm-level parameters to the other measured firm-level outcomes. We account for the estimation error $([\alpha_j \;\; \bar{u}_j\eta \;\; \phi_j \;\; \gamma_j \;\; \gamma_{2j}] - [\hat{\alpha}_j \;\; \bar{u}_j\hat{\eta} \;\; \hat{\phi}_j \;\; \hat{\gamma}_j \;\; \hat{\gamma}_{2j}])$ explicitly, but we do not add an additional measurement error. Thus, for example, we assert that the outcome p_j depends upon α_j and not upon $\alpha_j + \zeta_j$, where ζ_j is an independent measurement error.

Now, equation (3.45) can be re-expressed in a first order approximation around δ_j as

$$p_j = P_j'(\delta_j)\begin{bmatrix}\zeta\\ \rho\end{bmatrix} + \omega_j \tag{3.46}$$

where

$$\omega_j \equiv \left(\hat{\delta}_j - \delta_j\right)' \frac{\partial P_j'(\delta_j)}{\partial \delta_j}\begin{bmatrix}\zeta\\ \rho\end{bmatrix} + \zeta_j.$$

The variance of the regression error term for equation (3.46) consists of the component due to the estimation error in $\hat{P}_j$ plus the component due to ξ_j:

$$\operatorname{var}[\omega_j] \equiv \begin{bmatrix}\zeta' & \rho'\end{bmatrix} \frac{\partial P_j}{\partial \delta_j'} \operatorname{var}\left[\hat{\delta}_j\right] \frac{\partial P_j'}{\partial \delta_j}\begin{bmatrix}\zeta\\ \rho\end{bmatrix} + \operatorname{var}[\xi_j] \tag{3.47}$$

where the components of $\operatorname{var}[\hat{\delta}_j]$ are defined in the derivations above. We estimate equation (3.46) using generalized least squares based upon the error variance in equation (3.47).

4. DATA DESCRIPTION

In this section we describe the important institutions of the French labor market and compare some simple statistical models of wage determination in France and the United States. The wage regressions demonstrate that, even though French and American labor market institutions are quite different, there are strong similarities in the way compensation is related to labor market observables in the two countries. Next, we lay out the sample design of our French data and describe the process we used to create an analysis sample. Finally, we present all of the variable definitions. Summary statistics appear in the Data Appendix.

4.1. *The French Labor Market*

During the sample period (from the mid-seventies to the end of the eighties), the French labor market was characterized by stable employment, whereas over this period employment increased by 25% in the United States. GDP growth in both countries was more or less identical, implying faster productivity growth in France. In addition, the employment-population ratio in France shrank while it was growing in the U.S.; as a reference, employment-population ratios were the same in France and the United States in the mid-sixties. In particular, the employment-population ratio fell dramatically for young workers (below 25) as well as for older workers (above 55).[30] The prevailing view—challenged in Card,

[30] See Card, Lemieux, and Kramarz (1996), for a more detailed analysis of French labor market outcomes in comparison with those of the United States and Canada.

Kramarz, and Lemieux (1996)—is that wage rigidities, examples of which are presented in the following paragraphs, have destroyed jobs in France. Nevertheless, even though wage-setting institutions differ, wage setting outcomes in the two labor markets share many features.

French employment in the 1970s was characterized by centralized collective bargaining (convention collective de branche), in which different industrial sectors had collective agreements that were negotiated by groups of unions and employers associations, and these agreements were binding on the negotiating parties. The complete agreement was then typically extended to cover the entire industry (or region) by the Ministry of Labor and was thereby made binding on workers and firms that were not party to the original negotiation (see Margolis (1993)). More than 95% of the work force was covered by these collective bargaining agreements at the end of the 1980s, while union membership was approximately 10%. The collective agreements specified a set of minimum wages and wage progressions for the occupational categories covered by the negotiations (sometimes called a wage grid). Beginning in 1982, the “lois Auroux” (a set of revisions to the body of labor law named after the Minister of Labor at the time) required firms with at least 50 employees to negotiate firm-level collective agreements (accords d’entreprise). Although firms were explicitly *not* obligated to actually conclude an agreement, the percentage of the work force covered by firm-level agreements grew to over 30% by the mid-1980s (see Abowd and Kramarz (1993) and Cahuc and Kramarz (1997)). The law imposed that the firm-level agreements could only improve the conditions stated in the industrial agreement, a result being that, over time, the firm-level agreements have become more relevant for wage determination than the industry agreements.

Since 1951, French industry has also been subject to a national minimum wage (called the SMIC since the revisions to the relevant law in 1971) that is indexed to the rate of change in consumer prices and to the average blue-collar wage rate. Although more than 90% of French workers are covered by industrial agreements throughout our analysis period (1976–1987), the regular increases in the national minimum wage (in particular those driven by the indexation to the average blue-collar wage rate) outpaced contract renegotiations, and the lowest rungs on the pay scales in most industry contracts for most occupations ended up below the national minimum in 1985. When this occurs, it is the national minimum wage, and not the collectively bargained wage, that binds.

Even though the French institutional arrangements seem to differ widely from those prevailing in the United States, wage-setting outcomes in the two countries share many features. For instance, manufacturing operative wages, when measured in purchasing power parity, are not very different (see Abowd and Bognanno (1995)). However, the ratio of the minimum wage to the average wage fell sharply in the U.S. while it rose modestly in France during the eighties (see Card, Kramarz, and Lemieux (1996)). Roughly 7% of French employed young workers (30 years old and under), and 6% of American employed young workers are paid at the minimum during the same period (see Abowd, Kramarz, Lemieux, and Margolis (forthcoming)). Even though total labor costs at the minimum wage are higher in France than in the U.S. due to employee- and employer-paid payroll taxes and other nonwage compensation costs, a 1% increase in the minimum wage induces roughly a 2% decrease in employment of young people in both countries (Abowd, Kramarz, Lemieux, and Margolis (forthcoming)).

To further assess potential differences in wage setting, we ran two simple wage regressions using comparable household surveys (the Enquête Emploi for France and the Current Population Survey for the U.S.).[31] Table II presents our estimation results. Our models show that the same set of regressors has more or less the same explanatory power for wages in both the French and American data (roughly 37% for men in both countries, 32% for women in France and 24% in the U.S.). Returns to one additional year of education are 6.1% for men and 7.2% for women in the U.S. while they are 7.7% for men and 8.8% for women in France, with the difference between the sexes being identical. Returns to experience differ slightly, with the curvature of the quartic in experience implying a more hump-shaped profile in the U.S. Finally, the gender wage gap in the initial year is roughly equal in both countries, although it decreases over the sample period in the U.S. and is basically stable in France during the eighties.

Other examples of such similarities in wage-setting outcomes abound. Card, Kramarz, and Lemieux (1996) have shown that the fraction of workers using computers is roughly the same in the two countries. Furthermore, returns to new technologies, and in particular computer use, are identical in the two countries. Estimates in Krueger (1993), in Entorf and Kramarz (1997), or Entorf, Gollac, and Kramarz (forthcoming) show that computer users are better compensated than nonusers by the same amount (15%). Krueger and Summers (1987) also show that inter-industry wage differentials in France are highly correlated with American inter-industry wage differentials.

4.2. *Description of the DAS*

Our main data source is the "Déclarations Annuelles des Salaires" (DAS), a large-scale administrative database of matched employer-employee information collected by INSEE (Institut National de la Statistique et des Etudes Economiques) and maintained in the Division des Revenus. The data are based upon mandatory employer reports of the gross earnings of each employee subject to French payroll taxes. These taxes apply to all "declared" employees and to all self-employed persons, essentially all employed persons in the economy.

The Division des Revenus prepares an extract of the DAS for scientific analysis, covering all individuals employed in French enterprises who were born in October of even-numbered years, with civil servants excluded.[32] Our extract

[31] Similar results are also found using cross-sections of matched worker-firm data for the two countries (see Abowd, Kramarz, Margolis, and Troske (1998)).

[32] Meron (1988) shows that individuals employed in the civil service move almost exclusively to other positions within the civil service. Thus, the exclusion of civil servants should not affect our estimation of a worker's market wage equation. Employees of the state-owned firms are present in our sample, however.

TABLE II

COMPARISON OF LEAST SQUARES ESTIMATES OF WAGE DETERMINATION IN FRANCE AND THE UNITED STATES 1982–1987

	France				*United States*			
	Men		Women		Men		Women	
Variable	Mean	OLS Results	Mean	OLS Results	Mean	OLS Results	Mean	OLS Results
Intercept	1.000	1.365	1.000	1.163	1.000	0.534	1.000	0.380
	(0.000)	(6.746E-3)	(0.000)	(8.190E-3)	(0.000)	(5.614E-3)	(0.000)	(5.679E-3)
Years of Education	10.726	0.077	11.325	0.088	11.880	0.061	12.300	0.072
	(3.659)	(2.848E-4)	(3.267)	(3.998E-4)	(2.391)	(3.521E-4)	(2.149)	(3.712E-4)
Experience	20.722	0.058	19.048	0.060	15.894	0.112	16.036	0.082
	(12.222)	(1.228E-3)	(12.163)	(1.435E-3)	(12.311)	(9.219E-4)	(12.323)	(8.899E-4)
Experience^2/100	5.788	−0.104	5.108	−0.188	4.042	−0.506	4.090	−0.436
	(5.875)	(9.095E-3)	(5.533)	(1.133E-2)	(5.251)	(8.487E-3)	(5.077)	(8.482E-3)
Experience^3/1,000	18.915	−0.007	16.207	0.018	12.611	0.102	12.544	0.093
	(26.344)	(2.554E-3)	(23.730)	(3.355E-3)	(21.849)	(2.811E-3)	(20.514)	(2.898E-3)
Experience^4/10,000	68.009	0.002	56.700	0.001	43.914	−0.007	42.506	−0.007
	(120.682)	(2.397E-4)	(103.687)	(3.316E-4)	(92.920)	(3.037E-4)	(84.740)	(3.232E-4)
1982	0.175	0.036	0.167	0.027	0.163	0.072	0.160	0.019
	(0.380)	(3.001E-3)	(0.373)	(3.784E-3)	(0.370)	(2.715E-3)	(36.707)	(2.596E-3)
1983	0.170	0.018	0.166	0.006	0.162	0.049	0.160	0.015
	(0.375)	(3.020E-3)	(0.373)	(3.783E-3)	(0.369)	(2.707E-3)	(0.367)	(2.579E-3)
1984	0.166	0.019	0.164	0.012	0.164	0.032	0.162	0.000
	(0.372)	(3.033E-3)	(0.371)	(3.793E-3)	(0.370)	(2.679E-3)	(0.368)	(2.557E-3)
1985	0.165	0.005	0.165	−0.001	0.167	0.018	0.166	−0.002
	(0.371)	(3.040E-3)	(0.371)	(3.785E-3)	(0.373)	(2.658E-3)	(0.372)	(2.534E-3)
1986	0.162	0.023	0.168	0.018	0.174	0.015	0.175	0.004
	(0.369)	(3.051E-3)	(0.374)	(3.767E-3)	(0.379)	(2.601E-3)	(0.380)	(2.479E-3)
Paris Region	0.210	0.168	0.240	0.158	—	—	—	—
	(0.407)	(2.147E-3)	(0.427)	(2.567E-3)	—	—	—	—
Northeast U.S.	—	—	—	—	0.210	−0.046	0.217	−0.057
	—	—	—	—	(0.408)	(2.496E-3)	(0.412)	(2.393E-3)
Midwest U.S.	—	—	—	—	0.263	−0.039	0.273	−0.088
	—	—	—	—	(0.440)	(2.309E-3)	(0.446)	(2.222E-3)
Southern U.S.	—	—	—	—	0.296	−0.143	0.289	−0.128
	—	—	—	—	(0.457)	(2.206E-3)	(0.453)	(2.151E-3)
Observations	165,036	165,036	126,320	126,320	259,297	259,297	259,266	259,266
Adjusted R^2		0.3866		0.3254		0.3626		0.2428

Sources: Enquête Emploi (1982–1987) for France and NBER outgoing rotation group CPS extracts (1982–1987) for the United States. *Notes:* Standard Deviations/Errors in Parentheses. Both regressions included only individuals between 16 and 60 years old, inclusive. Both regressions used the sample weights. Experience is measured as (age) − (age at the end of schooling) in France and (age) − (years of schooling) − 5 in the United States.

runs from 1976 through 1987, with 1981 and 1983 excluded because the underlying administrative data were not sampled in those years. The initial data set contained 7,416,422 observations. Each observation corresponded to a unique individual-year-establishment combination. An observation in this initial DAS file includes an identifier that corresponds to the employee (called ID below), an identifier that corresponds to the establishment (SIRET) and an identifier that corresponds to the parent enterprise of the establishment (SIREN). For each observation, we have information on the number of days during the calendar year the individual worked in the establishment and the full-time/part-time status of the employee. For each observation corresponding to an individual-year-establishment, in addition to the variables listed above, we have information on the individual's sex, date and place of birth, occupation, total net nominal earnings during the year and annualized gross nominal earnings during the year for the individual, as well as the location and industry of the employing establishment.

4.3. *Observation Selection, Variable Creation and Missing Data Imputation*

4.3.1. *Aggregation across establishments*

The creation of the analysis data set involved the selection of desired individuals, the aggregation of establishment-level data to the enterprise level, and the construction of the variables of interest from the variables already in the data set. We selected only full-time employees (sample reduced to 5,966,620 observations). We then created a single observation for each ID-year-SIREN combination by aggregating within ID and year over SIRETs in the same SIREN. For each ID-year-SIREN, we summed total net nominal earnings and total days worked over all SIRETs. We assigned to the observation the occupation, location, and industry that corresponded to the establishment in which the individual worked the largest number of days during the year. This reduced the number of observations to 5,965,256. We then selected the enterprise at which the individual had worked the largest number of days during that year (sample reduced to 5,497,287 observations). The aggregation of total number of days worked across all establishments occasionally yielded observations for which the total number of days worked was greater than 360 (the maximum permitted). In these cases, we truncated days worked at 360. We then calculated an annualized net nominal earnings for the ID-year-SIREN combination. We eliminated all years of data for individuals who were younger than 15 years old or older than 65 years old at the date of their first appearance in the data set (sample reduced to 5,325,413 observations).

4.3.2. *Total compensation costs*

The dependent variable in our wage rate analysis is the logarithm of real annualized total compensation cost for the employee. To convert the annualized net nominal earnings to total compensation costs, we used the tax rules and computer programs provided by the Division des Revenus at INSEE (Lhéritier, internal, undated INSEE communication) to compute both the employee and employer share of all mandatory payroll taxes (cotisations et charges salariales: employé et employeur). Annualized total compensation cost is defined as the sum of annualized net nominal earnings, annualized employee payroll taxes, and annualized employer payroll taxes. Nominal values were deflated by the consumer price index to get real annualized net earnings, and real annualized total compensation cost. We eliminated 61 observations with zero values for annualized total compensation cost (remaining sample 5,325,352).

4.3.3. *Education and school-leaving age*

Our initial DAS file did not contain education information. We used supplementary information from the permanent demographic sample (Echantillon Démographique Permanent, EDP) available for 10% of the DAS, to impute the level of education for all remaining individuals in the DAS. The EDP includes information on the highest degree obtained. There were 38 possible responses, including "no known degree." These responses were grouped into eight degree-level categories as shown in Data Appendix Table B1. Using these eight categories as the dependent variable and data available in the DAS, we ran separate ordered logits for men and women. We used the data corresponding to the earliest date that an individual appeared in our sample to estimate these models. We used the same data and the estimated coefficients from these ordered logit models to impute the probability of obtaining each of the eight different aggregated degrees for the individuals in the DAS who were not part of the EDP. We used the actual value of the eight degree aggregates for the EDP sample members. Thus, a random 10% sample of the DAS individuals have true education and the remaining 90% have the probability of obtaining each of the eight degree aggregates. EDP sample statistics for the men are in Data Appendix Table B2, and those for the women are in Table B3. The estimated logit equations are in Table B4 for men and Table B5 for women.[33]

To calculate school-leaving age we used Table 14 in CEREQ-DEP-INSEE (1990), which provides the average age of termination for each French diploma separately for men and women in 1986. Using the probability of each degree category and the average school-leaving age for degrees in that category (the ages were fairly homogeneous within categories), we calculated expected school-leaving age.

4.3.4. *Total labor market experience*

For the first year in which individuals appear in the sample, we calculate potential labor market experience as age at the beginning of the year less our estimate of school-leaving age. In all subsequent years, total labor market experience is accumulated using the individual's realized labor force history. Our algorithm was the same for both labor force experience and seniority. It

[33] We considered, and rejected, the possibility of using a Rubin (1987) style multiple imputation algorithm for the missing schooling variable for the following reasons: (1) since schooling does not time-vary, and since our conclusions are completely unaffected by whether we remove a schooling effect from the person effect or not (θ as compared with α), we did not want to bear the computational burden associated with these methods for such a small return; (2) the schooling variable is substituting for occupational category, a more common control variable in French earnings equations because of the educational qualifications that define the occupational categories, in our models in order to make the analysis more comparable to the vast American literature that uses schooling rather than occupation and defines person effects with a schooling effect removed; (3) conditioning the imputation on the observed value of the compensation variable, as these methods require, would focus attention on the imputation procedure and detract from our main focus.

accounts for the holes caused by the fact that the administrative data were not available for 1981 and 1983. See the Data Appendix for details.

4.3.5. *Job seniority*

Individuals fell into two categories with respect to the calculation of job seniority (employer-specific experience): those for whom the first year of observation was 1976 and those who first appeared after 1976. For those individuals whose first observation was in 1976, we estimated the expected length of the in-progress employment spell by regression analysis using a supplementary survey, the 1978 Salary Structure Survey (Enquête sur la Structure des Salaires, ESS). In this survey, respondent establishments provided information on seniority (in 1978), occupation, date of birth, industry, and work location for a scientific sample of their employees. Using the ESS information, we estimated separate regressions for men and women to predict seniority for in-progress spells in 1976. The coefficients from these regressions were used to calculate expected job seniority in 1976 for DAS individuals whose first observation was in 1976. The dependent variable in the supplementary ESS regressions was current seniority with the employer and the explanatory variables were date of birth, occupation (1-digit), region of employment (metropolitan Paris), and industry (NAP 100, approximately 2-digit). Results of the seniority regressions are shown in equations (8.1) for men and (8.2) for women in the Data Appendix. We used the results of these seniority regressions to impute levels of job seniority in 1976 for the left-censored DAS individuals first observed in 1976. Details are provided in the Data Appendix.

4.3.6. *Elimination of outliers*

After calculating all of the individual-level variables, we eliminated observations for which the log of the annualized real total compensation cost was more than five standard deviations away from its predicted value based on a linear regression model with independent variables: sex, region of France, experience and education (see equation (8.4) in the Data Appendix). This gives us the analysis sample of 5,305,108 observations.

Table B7 in the Data Appendix shows the basic summary statistics, by sex, for the individual-level data. The usable sample consists of 3,434,530 observations on 711,518 men and 1,870,578 usable observations on 454,787 women. The basic individual-level variables are: sex, labor force experience, region of France, education, and seniority. Note that about 30% of the sample has no known educational attainment. For 74% of the individuals, there are enough observations in the sample to permit estimation of a distinct firm-effect.[34]

[34] The individuals from firms with fewer than 10 observations in the sample were pooled and a single firm-level regression was used to estimate their firm effects.

4.4. *Construction of the Firm-Level Data*

4.4.1. *DAS-based firm-level averages*

For our firm-level analyses we calculated the aggregates $\hat{\alpha}_j$, $\bar{u}_j\hat{\eta}$ and their respective sampling variances based on the $\hat{\alpha}_i$ and $u_i\hat{\eta}$ estimated according to the conditional estimation methods laid out in Section 3 above. The estimated parameters $\hat{\phi}_j$, $\hat{\gamma}_j$, and $\hat{\gamma}_{2j}$ have unique values for a given enterprise, by construction. In cases where any one of the following three conditions failed: $-3 \leq \hat{\phi}_j \leq 3$ or $-2 \leq \hat{\gamma}_j \leq 2$ or $-2 \leq \hat{\gamma}_j + \hat{\gamma}_{2j} \leq 2$, we set $\hat{\phi}_j$, $\hat{\gamma}_j$, and $\hat{\gamma}_{2j}$ equal to the values estimated in the pooled model for the firm effects for all firms with 10 or fewer observations.

4.4.2. *Other firm-level data sources*

The primary source of our firm-level data is the INSEE (1989, 1990a–c) enterprise sample (Echantillon d'Entreprises, EE), a probability sample of French firms (synonymous with enterprises for our purposes). The EE data set provides the sampling frame for the firm-based part of this paper. The universe for the sample is the annual report on profitability and employment by enterprises (Bénéfices Industriels et Commerciaux, BIC) and the annual survey of enterprises (Enquête Annuelle d'Entreprises, EAE). To construct the EE, firms with more than 500 employees were sampled from the BIC with probability 1; firms with 50 to 499 employees were sampled from the BIC with probabilities ranging from 1/4 to 1/2 depending upon the industry, and smaller firms were sampled from the BIC with probability 1/30. All firms responding to the BIC were at risk to be sampled exactly once. Hence, the EE is dynamically representative of French enterprises in all sectors except the public administration sector. We use the sampling weight (non-time-varying) and the variables described below, averaged over the period 1978 to 1988 for all available years, from the EE.

4.4.3. *Firm-level employment and capital stock*

The measure of employment, in thousands of workers, is full-time employment in an enterprise as of December 31 (prior to 1984) and the annual average full-time employment (1984 and later) as found in the BIC. We took the mean of this variable over all years that the firm appeared in the sample.

Total capital in the enterprise is defined as the sum of debt (dettes) and owners' equity (fonds propres d'entreprise). Our capital measure is equal to total assets (actif total) in French accounting systems. The information was taken directly from the BIC for every firm-year. We deflated the capital stock using an industry-specific, annual index of the price of capital from the INSEE macroeconomic time series data (Banque de Données Macroéconomiques, BDM). Our measure of real total capital is defined as total assets divided by the industry-specific price index of physical capital (in millions of 1980 FF), averaged over all available years for the firm. The capital labor ratio is defined as real total capital divided by total full-time employment. We also averaged this variable over all available years for the firm.

4.4.4. *Real operating income per unit of capital*

We used the BIC to obtain the operating income (excédent brut d'exploitation, EBE), for each firm in each year that it appeared in the firm sample. The formula used to calculate the EBE is shown in equation (8.6) in the Data Appendix. The EBE was deflated by the value added price index (prix de valeur ajoutée) also found in the BDM, to yield real operating income (in millions of 1980 FF). Real operating income was divided by real total capital to yield real operating income per unit of capital, stated as a proportion. We also took the mean of this variable over all available years for the firm.

4.4.5. *Real value added inclusive of labor costs*

To calculate the real value added inclusive of labor costs (valeur ajoutée réelle brute au coût des facteurs), we divided the employer's compensation costs (frais de personnel) in the BIC (thousands of FF) by the consumer price index (indice des prix à la consommation) from the BDM to yield the employer's real compensation cost (in millions of 1980 FF). The result was added to real operating income, as defined above in Section 4.4.4, to yield the real value added inclusive of labor costs (in millions of 1980 FF). Real value added inclusive of labor costs was divided by total employment to yield real value added inclusive of labor costs per worker (in thousands of 1980 FF). We then took the mean of this variable over all of the years that the firm appeared in the EE sample.

4.4.6. *Employment structure*

The variables concerning the occupational structure of employment (proportion of engineers, technicians and managers in the work force and proportion of skilled workers) were created from the employment structure survey (Enquête sur la Structure des Emplois, ESE), which is an annual administrative data base of the detailed (4-digit) occupational structure of all establishments with more than 20 employees. The occupational structure of the firm, measured in the ESE, was merged with the EE using the firm identifier and the survey year. Engineers, technicians and managers were coded using the simplified occupation classifications (1-digit equivalents) for individuals in categories 30 and 40. The proportion of skilled workers in the work force was calculated from the ESE using the simplified occupation classification for individuals in categories 50 and 61. Both variables were expressed as a ratio to total employment and averaged over all the available firm-years. The omitted category is unskilled workers, which would include all other codes.

5. ESTIMATION RESULTS

5.1. *Overview*

We present our statistical results in three main parts. Tables III–VI present detailed results from the analysis of the matched employer-employee microdata. Table III shows the regression coefficients for men and women from all of the estimation methods described above as well as results from standard specifications based upon incomplete parameterizations of equation (2.2). Table IV presents summary statistics for men and women for all of the components of compensation in our complete model and for the two estimation methods upon which we focus most of our subsequent attention. Table V is a diagnostic table of the correlations among the same components of compensation when we vary the method of estimation. Table VI is a table of correlations among all the components of compensation for our two chosen estimation methods. Tables VII and VIII present our statistical analysis of the inter-industry wage differential and the firm-size wage effect, respectively. Tables IX–XII present the results of analyses conducted at the firm level. Table IX shows summary statistics for the firm-level variables, including those we created from the matched microdata. Table X presents the results of our analysis of firm-level profitability and productivity. Table XI presents our analysis of firm-level factors of production and compensation components. Table XII presents the results of a survival analysis using the firm-level data.

5.2. *Results from the Estimation of the General Compensation Equation*

5.2.1. *Specification of the compensation equation*

Table III presents a summary of the estimates of β, the coefficients on the time-varying individual characteristics, for our consistent estimation method, our conditional estimation method with persons first, and ordinary least squares under a variety of different assumptions about the included person and firm effects, separately for men and women. The results labelled "Consistent Method Person & Firm Effects" were calculated according to equation (3.9). The results labelled "Conditional Method Persons First" were calculated using the formula found in equation (3.17). The β coefficients obtained by the order-dependent method with persons first, conditional on Z, and those obtained by the order-independent method, conditional on Z, are mathematically identical. The column labelled "Least Squares No Person/Firm Effects" presents the estimates obtained when we set all person effects, $D\theta$, firm effects, $F\psi$, and conditioning effects, $Z\lambda$, jointly to zero. Next, in the column labelled "Within Persons No Firm Effects," we present results obtained when we retain the person effects, $D\theta$, but set all firm effects, $F\psi$, and conditioning effects, $Z\lambda$, jointly to zero. In the column labelled "Within Persons Limited Firm Effects," we present β coefficients estimated when we retain all person effects, $D\theta$, choose a set of effects Z equal to the columns of F corresponding to the 115 largest employers

TABLE III

ESTIMATES OF THE EFFECTS OF LABOR FORCE EXPERIENCE, REGION, YEAR, EDUCATION, INDIVIDUALS, AND FIRMS ON THE LOG OF REAL TOTAL ANNUAL COMPENSATION COSTS INDIVIDUAL DATA BY SEX FOR 1976 TO 1987

Variable	*Consistent Method Person & Firm Effects* Parameter Estimate	Standard Error	*Conditional Method Persons First* Parameter Estimate	Standard Error	*Least Squares No Person/ Firm Effects* Parameter Estimate	Standard Error	*Within Persons No Firm Effects* Parameter Estimate	Standard Error	*Within Persons Limited Firm Effects*[a] Parameter Estimate	Standard Error	*Within Firms No Person Effects* Parameter Estimate	Standard Error	*Within Industry No Person Effects* Parameter Estimate	Standard Error	*Within Firm Size No Person Effects* Parameter Estimate	Standard Error
Men																
Total Labor Force Experience	0.0586	(0.0015)	0.0687	(0.0004)	0.0542	(0.0003)	0.0695	(0.0004)	0.0685	(0.0004)	0.0448	(0.0003)	0.0521	(0.0003)	0.0507	(0.0003)
(Labor Force Experience)^2/100	−0.3432	(0.0072)	−0.4415	(0.0027)	−0.2280	(0.0030)	−0.4543	(0.0029)	−0.4446	(0.0030)	−0.1584	(0.0027)	−0.2115	(0.0030)	−0.2047	(0.0029)
(Labor Force Experience)^3/1,000	0.0734	(0.0025)	0.1053	(0.0010)	0.0503	(0.0010)	0.1100	(0.0010)	0.1074	(0.0010)	0.0298	(0.0009)	0.0452	(0.0010)	0.0441	(0.0010)
(Labor Force Experience)^4/10,000	−0.0057	(00003)	−0.0093	(0.0001)	−0.0046	(0.0001)	−0.0099	(0.0001)	−0.0096	(0.0001)	−0.0025	(0.0001)	−0.0041	(0.0001)	−0.0040	(0.0001)
Lives in Ile-de-France	0.0051	(0.0017)	0.0832	(0.0010)	0.1398	(0.0005)	0.0819	(0.0011)	0.0805	(0.0011)	0.1117	(0.0007)	0.1316	(0.0006)	0.1314	(0.0005)
Year 1977	0.0245	(0.0009)	0.0251	(0.0007)	0.0343	(0.0010)	0.0252	(0.0007)	0.0248	(0.0007)	0.0182	(0.0009)	0.0275	(0.0010)	0.0296	(0.0010)
Year 1978	0.0643	(0.0017)	0.0605	(0.0008)	0.0645	(0.0010)	0.0609	(0.0008)	0.0598	(0.0008)	0.0463	(0.0009)	0.0560	(0.0010)	0.0581	(0.0010)
Year 1979	0.0887	(0.0024)	0.0879	(0.0009)	0.0841	(0.0010)	0.0883	(0.0010)	0.0873	(0.0010)	0.0598	(0.0009)	0.0755	(0.0010)	0.0774	(0.0010)
Year 1980	0.1081	(0.0031)	0.1030	(0.0011)	0.0899	(0.0010)	0.1033	(0.0012)	0.1024	(0.0012)	0.0644	(0.0009)	0.0803	(0.0010)	0.0841	(0.0010)
Year 1982	0.1473	(0.0035)	0.1441	(0.0014)	0.1137	(0.0010)	0.1447	(0.0016)	0.1434	(0.0016)	0.0809	(0.0009)	0.1043	(0.0010)	0.1091	(0.0010)
Year 1984	0.1872	(0.0041)	0.1911	(0.0018)	0.1441	(0.0010)	0.1919	(0.0020)	0.1903	(0.0020)	0.1009	(0.0009)	0.1316	(0.0011)	0.1386	(0.0011)
Year 1985	0.2044	(0.0047)	0.2173	(0.0020)	0.1662	(0.0011)	0.2179	(0.0022)	0.2162	(0.0022)	0.1146	(0.0009)	0.1516	(0.0011)	0.1612	(0.0011)
Year 1986	0.2366	(0.0053)	0.2529	(0.0022)	0.1841	(0.0010)	0.2535	(0.0024)	0.2517	(0.0024)	0.1315	(0.0009)	0.1690	(0.0011)	0.1813	(0.0011)
Year 1987	0.2499	(0.0060)	0.2763	(0.0024)	0.1954	(0.0010)	0.2768	(0.0026)	0.2749	(0.0026)	0.1401	(0.0009)	0.1808	(0.0011)	0.1948	(0.0011)
Women																
Total Labor Force Experience	0.0144	(0.0016)	0.0290	(0.0005)	0.0326	(0.0004)	0.0308	(0.0006)	0.0298	(0.0005)	0.0224	(0.0004)	0.0603	(0.0004)	0.0286	(0.0004)
(Labor Force Experience)^2/100	−0.1063	(0.0091)	−0.1728	(0.0036)	−0.1117	(0.0038)	−0.1771	(0.0041)	−0.1729	(0.0040)	−0.0318	(0.0035)	−0.3525	(0.0039)	−0.0816	(0.0038)

(Labor Force Experience)^3/1,000	0.0184 (0.0032)	0.0379 (0.0013)	0.0183 (0.0013)	0.0391 (0.0014)	0.0381 (0.0014)	−0.0053 (0.0012)	0.0942 (0.0013)	0.0103 (0.0013)
(Labor Force Experience)^4/10,000	−0.0009 (0.0004)	−0.0031 (0.0001)	−0.0013 (0.0001)	−0.0031 (0.0002)	−0.0031 (0.0002)	0.0011 (0.0001)	−0.0091 (0.0001)	−0.0005 (0.0001)
Lives in Ile-de-France	0.0042 (0.0027)	0.0795 (0.0016)	0.1576 (0.0007)	0.0794 (0.0018)	0.0809 (0.0017)	0.1218 (0.0009)	0.1434 (0.0007)	0.1470 (0.0007)
Year 1977	0.0300 (0.0011)	0.0271 (0.0009)	0.0527 (0.0014)	0.0250 (0.0011)	0.0255 (0.0010)	0.0348 (0.0012)	0.1361 (0.0015)	0.0495 (0.0014)
Year 1978	0.0762 (0.0019)	0.0724 (0.0010)	0.1053 (0.0014)	0.0688 (0.0012)	0.0695 (0.0011)	0.0798 (0.0012)	0.1889 (0.0015)	0.1003 (0.0014)
Year 1979	0.1102 (0.0026)	0.1052 (0.0012)	0.1353 (0.0014)	0.1003 (0.0014)	0.1015 (0.0013)	0.1044 (0.0012)	0.2179 (0.0014)	0.1303 (0.0014)
Year 1980	0.1329 (0.0033)	0.1227 (0.0014)	0.1445 (0.0014)	0.1169 (0.0016)	0.1182 (0.0015)	0.1148 (0.0012)	0.2285 (0.0014)	0.1408 (0.0014)
Year 1982	0.1830 (0.0039)	0.1704 (0.0018)	0.1758 (0.0014)	0.1627 (0.0020)	0.1640 (0.0019)	0.1406 (0.0012)	0.2600 (0.0015)	0.1742 (0.0015)
Year 1984	0.2233 (0.0047)	0.2188 (0.0022)	0.2231 (0.0014)	0.2094 (0.0025)	0.2109 (0.0024)	0.1719 (0.0013)	0.3021 (0.0015)	0.2200 (0.0015)
Year 1985	0.2361 (0.0053)	0.2377 (0.0024)	0.2392 (0.0014)	0.2277 (0.0027)	0.2292 (0.0026)	0.1782 (0.0013)	0.3163 (0.0015)	0.2360 (0.0015)
Year 1986	0.2644 (0.0059)	0.2686 (0.0026)	0.2559 (0.0014)	0.2577 (0.0030)	0.2594 (0.0029)	0.1945 (0.0013)	0.3340 (0.0015)	0.2549 (0.0015)
Year 1987	0.2756 (0.0066)	0.2886 (0.0028)	0.2615 (0.0014)	0.2767 (0.0033)	0.2787 (0.0031)	0.1995 (0.0013)	0.3414 (0.0015)	0.2630 (0.0015)
Pooled								
Sample Size	5,305,108	5,305,108	5,305,108	5,305,108	5,305,108	5,305,108	5,109,008[b]	5,305,108
Coefficient Degrees of Freedom (β)	28	28	44	30	30	45	129	72
Coefficient Degrees of Freedom (λ)		48[c]						
Individual Degrees of Freedom (θ)	2,011,864	1,166,305		1,166,305	1,166,305			
Firm Degrees of Freedom (ψ)	521,182				229	521,182		
Error Degrees of Freedom (ε)	2,772,034	4,138,727	5,305,064	4,138,773	4,138,544	4,783,881	5,108,879	5,305,036
Root Mean Squared Error	0.2828	0.2732	0.4223	0.2737	0.2733	0.3577	0.4204	0.4179
R^2	0.8364	0.7720	0.3017	0.7711	0.7720	0.5482	0.3389	0.3164
Specification Test	na	3806.5	7543.6	3410.5	3400.1	6642.0	na	na

Notes: (a) Includes firm effects (intercept and seniority slope) for the 115 largest firms in the sample. (b) Total degrees of freedom differ from other columns because of missing industry data. (c) See Data Appendix Table B6 for summary statistics, coefficients and standard errors corresponding to these variables.

TABLE IV

DESCRIPTIVE STATISTICS FOR COMPONENTS OF LOG REAL TOTAL ANNUAL COMPENSATION BY SEX FOR 1976 TO 1987

	Men		*Women*	
Variable Definition	Mean	Std. Dev.	Mean	Std. Dev.
Log (Real Annual Compensation, 1980 FF)	4.3442	0.5187	4.0984	0.4801
Order-Independent Method				
$x\beta$, Predicted Effect of x Variables	0.3890	0.1489	0.2849	0.1144
θ, Individual Effect Including Education	3.9552	0.4475	3.8135	0.3930
α, Individual Effect (Unobserved Factors)	0.0000	0.4051	0.0000	0.3771
$u\eta$, Individual Effect of Education and Sex	3.9552	0.1902	3.6893	0.1107
ψ, Firm Effect (Intercept and Slope)	−0.0363	0.4642	0.0665	0.5116
ϕ, Firm Effect Intercept	−0.1367	0.4532	−0.0235	0.4967
γ, Firm Effect Slope	0.0149	0.0503	0.0172	0.0531
Order-Dependent Method: Persons First				
$x\beta$, Predicted Effect of x Variables	0.4261	0.1383	0.3234	0.1120
θ, Individual Effect Including Education	3.9160	0.4387	3.7776	0.3843
α, Individual Effect (Unobserved Factors)	0.0000	0.3947	0.0000	0.3639
$u\eta$, Individual Effect of Education and Sex	3.9160	0.1915	3.7776	0.1238
ψ, Firm Effect (Intercept and Slope)	0.0028	0.0685	−0.0039	0.0566
ϕ, Firm Effect Intercept	0.0031	0.1044	−0.0072	0.0969
γ, Firm Effect Slope	−3.37*e*-05	0.0335	8.28*e*-04	0.0326
γ_2, Firm Effect Slope Change at 10 years	−5.36*e*-04	0.0542	−1.64*e*-03	0.0574
Other Estimation Methods				
Seniority coefficient, Least Squares (Standard Error)	0.0118	(0.0001)	0.0141	(0.0001)
Seniority coefficient, within Persons (Standard Error)	0.0033	(0.0001)	0.0024	(0.0001)
Seniority coefficient, within Firms (Standard Error)	0.0078	(0.0001)	0.0102	(0.0001)
Seniority coefficient, within Industry (Standard Error)	0.0090	(0.0001)	0.0121	(0.0001)
Seniority coefficient, within Size Class (Standard Error)	0.0097	(0.0001)	0.0126	(0.0001)
γ, Firm Effect Slope, 115 largest firms	0.0013	0.0065	0.0014	0.0076
γ, Firm Effect Slope, Consistent estimates	0.0116	0.0342	0.0138	0.0352

Notes: Seniority coefficients with standard errors were estimated in the same models reported in Table III. All other statistics are the means and standard deviations based upon the sample of 5,305,108 observations except for the Firm Effect Slope in the 115 largest firms, which are statistics based on 695,077 observations.

in our data (firm-specific intercepts and seniority slopes), and set all remaining firm effects, $F\psi$, to a single common effect. Thus, this column shows estimates of a model in which 695 thousand of our 5.3 million observations have separate, firm effects (firm-specific intercepts and seniority slopes) and all remaining observations are pooled into a single artificial "firm," which had its own

TABLE V

CORRELATIONS AMONG THE COMPONENTS OF PERSON AND FIRM HETEROGENEITY AS ESTIMATED BY THE ORDER-INDEPENDENT, ORDER-DEPENDENT, FULL LEAST SQUARES ON THE 115 LARGEST FIRMS, AND CONSISTENT METHODS

Source of Estimate of the Indicated Effect	Parameter Name	Simple Correlation with: Order-Independent Estimates		Order-Dependent Estimates			Full Least Squares Estimates on the 115 Largest Firms		Consistent Estimates
		Persons First							
Firm Effects		ϕ	γ	ϕ	γ	γ_2	ϕ	γ	γ
Order-Independent	ϕ	1.0000	−0.0718	0.1553	−0.0837	0.0188	0.0888	0.2800	0.0361
Estimates	γ	−0.0718	1.0000	−0.2202	0.5300	−0.0077	−0.3276	0.3126	0.0907
Order-Dependent	ϕ	0.1553	−0.2202	1.0000	−0.5625	0.2562	0.6659	−0.0231	−0.1810
Estimates	γ	−0.0837	0.5300	−0.5625	1.0000	−0.2094	−0.6580	0.2739	0.1358
(Persons First)	γ_2	0.0188	−0.0077	0.2562	−0.2094	1.0000	0.5492	0.0293	−0.0126
Full Least Squares	ϕ	0.0888	−0.3276	0.6659	−0.6580	0.5492	1.0000	−0.1841	−0.1964
Estimates Using the 115 Largest Firms	γ	0.2800	0.3126	−0.0231	0.2739	0.0293	−0.1841	1.0000	0.5106
Consistent Estimates	γ	0.0361	0.0907	−0.1810	0.1358	−0.0126	−0.1964	0.5106	1.0000
		Firms First							
Person Effects		α		α			α		
Order-Independent Estimates	α	1.0000		0.5833			0.9896		
Order-Dependent Estimates (Firms First)	α	0.5833		1.0000			0.5983		
Full Least Squares Estimates Using the 115 Largest Firms	α	0.9896		0.5983			1.0000		

Notes: $N = 5{,}305{,}108$, except for Full Least Squares Estimates Using the 115 Largest Firms where $N = 695{,}077$.
Source: Authors' calculations based on the DAS.

intercept and seniority slope. Table III also shows, in the column labelled "Within Firms No Person Effects," the results obtained when we estimate a model where we retain all firm effects, $F\psi$ (intercepts only), and set all person effects, $D\theta$, and all conditioning effects, $Z\lambda$, to zero. The last two columns present results obtained from estimating a model with person effects and firm effects jointly set to zero, and with the functions Z of the form $Z = FA$, where A generates 84 industry effects and is as defined in Section 2.1 ("Within Industry No Person Effects"), and of the form $Z = FS$, with the matrix S generating 25 firm-size classes based on the firm sizes constructed as described in Section 3.6 ("Within Firm Size No Person Effects").

All of the estimation methods that include person effects (consistent method, conditional method with persons first or order independent, within persons without firm effects, and within persons with limited firm effects) are able to explain a similar fraction of the variance—between 77% and 83%. In contrast, all of the results that exclude person effects (ordinary least squares, within firms, within industries, and within firm-size categories) give results similar to the

TABLE VI

SUMMARY STATISTICS FOR THE DECOMPOSITION OF VARIANCE USING THE ORDER-INDEPENDENT AND THE ORDER-DEPENDENT CONDITIONAL METHODS FOR INDIVIDUAL DATA, BOTH SEXES, 1976–1987

Order-Independent Estimation Variable Description	Mean	Std. Dev.	Simple Correlation with: y	$x\beta$	θ	α	$u\eta$	ψ	ϕ	$s\gamma$	γ
y, Log (Real Annual Compensation, 1980 FF)	4.2575	0.5189	1.0000	0.2614	0.8962	0.8015	0.4011	0.2604	0.1603	0.2729	0.0333
$x\beta$, Predicted Effect of x Variables	0.3523	0.1464	0.2614	1.0000	−0.0445	−0.1243	0.1509	0.0697	0.0824	−0.0279	0.0300
θ, Individual Effect Including Education[a]	3.9052	0.4335	0.8962	−0.0445	1.0000	0.8964	0.4433	0.2965	0.1717	0.3384	0.0387
α, Individual Effect (Unobserved Factors)[a]	0.0000	0.3955	0.8015	−0.1243	0.8964	1.0000	0.0000	0.2640	0.1465	0.3178	0.0372
$u\eta$, Individual Effect of Education	3.9052	0.1776	0.4011	0.1509	0.4433	0.0000	1.0000	0.1349	0.0910	0.1209	0.0122
ψ, Firm Effect (Intercept and Slope)	0.0000	0.4839	0.2604	0.0697	0.2965	0.2640	0.1349	1.0000	0.9259	0.2537	0.0860
ϕ, Firm Effect Intercept	−0.0968	0.4721	0.1603	0.0824	0.1717	0.1465	0.0910	0.9259	1.0000	−0.1305	−0.0718
$s\gamma$, Firm Effect of Seniority	0.0968	0.1844	0.2729	−0.0279	0.3384	0.3178	0.1209	0.2537	−0.1305	1.0000	0.4094
γ, Firm Effect Slope	0.0157	0.0513	0.0333	0.0300	0.0387	0.0372	0.0122	0.0860	−0.0718	0.4094	1.0000

Order-Dependent Estimation: Persons First Variable Description	Mean	Std. Dev.	Simple Correlation with: y	$x\beta$	θ	α	$u\eta$	ψ	ϕ	$s\gamma + T(s)\gamma_2$	γ	γ_2
y, Log (Real Annual Compensation, 1980 FF)	4.2575	0.5189	1.0000	0.3271	0.9310	0.7331	0.4143	0.2131	0.1303	0.0053	0.0293	0.0276
$x\beta$, Predicted Effect of x Variables	0.3899	0.1386	0.3271	1.0000	0.0787	−0.0290	0.2211	0.0325	0.0350	−0.0157	−0.0148	0.0077
θ, Individual Effect Including Education[a]	3.8672	0.4255	0.9310	0.0787	1.0000	0.8842	0.4769	0.1079	0.0889	−0.0223	−0.0190	0.0225
α, Individual Effect (Unobserved Factors)[a]	0.0000	0.3841	0.7331	−0.0290	0.8842	1.0000	0.0000	0.0926	0.0828	−0.0263	−0.0202	0.0202
$u\eta$, Individual Effect of Education	3.8672	0.1831	0.4143	0.2211	0.4769	0.0000	1.0000	0.0473	0.0263	0.0041	−0.0006	0.0081
ψ, Firm Effect (Intercept and Slope)	0.0000	0.0647	0.2131	0.0325	0.1079	0.0926	0.0473	1.0000	0.4428	0.2089	−0.0909	0.0717
ϕ, Firm Effect Intercept	−0.0009	0.1019	0.1303	0.0350	0.0889	0.0828	0.0263	0.4428	1.0000	−0.7844	−0.5625	0.2562
$s\gamma + T(s)\gamma_2$, Firm Effect of Seniority	0.0009	0.0935	0.0053	−0.0157	−0.0223	−0.0263	0.0041	0.2089	−0.7844	1.0000	0.5507	−0.2298
γ, Firm Effect Slope	0.0003	0.0332	−0.0293	−0.0148	−0.0190	−0.0202	−0.0006	−0.0909	−0.5625	0.5507	1.0000	−0.2094
γ_2, Change in Firm Effect Slope	−0.0009	0.0553	0.0276	0.0077	0.0225	0.0202	0.0081	0.0717	0.2562	−0.2298	−0.2094	1.0000

Notes: (a) Correlations have been corrected for the sampling variance of the estimated effect.

TABLE VII

GENERALIZED LEAST SQUARES ESTIMATION BETWEEN INDUSTRY WAGE EFFECTS AND INDUSTRY AVERAGES OF FIRM-SPECIFIC COMPENSATION POLICIES

Independent Variable	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
		Based on Order-Independent Estimates				
Industry Average α	1.0390	(0.0023)	1.0053	(0.0022)		
Industry Average ψ	-0.0220	(0.0006)			0.0683	(0.0005)
Intercept	3.3023	(0.0019)	3.3031	(0.0019)	3.0935	(0.0018)
R^2	0.8487		0.8425		0.0682	
		Based on Order-Dependent Estimates: *Persons First*				
Industry Average α	0.8011	(0.0019)	0.8324	(0.0017)		
Industry Average ψ	0.2410	(0.0151)			-0.6659	(0.0150)
Intercept	3.1126	(0.0019)	3.1088	(0.0018)	3.0687	(0.0019)
R^2	0.9580		0.9213		0.2486	

Notes: The dependent variable is the 84 industry-effects estimated by least squares controlling for labor force experience (through quartic), seniority, region, year, education (eight categories) and sex (fully interacted). See Table III for the regression results. The independent variables are the industry averages for the indicated firm-specific compensation policy, adjusted for the same independent variables. The time period is 1976–1987.

TABLE VIII

GENERALIZED LEAST SQUARES ESTIMATES OF THE RELATION BETWEEN FIRM-SIZE WAGE EFFECTS AND FIRM-SIZE CATEGORY AVERAGES OF FIRM-SPECIFIC COMPENSATION POLICIES

Independent Variable	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
		Based on Order-Independent Estimates				
Firm-Size Category Average α	1.2222	(0.0043)	1.3245	(0.0041)		
Firm-Size Category Average ψ	0.2233	(0.0026)			0.4278	(0.0025)
Intercept	3.7397	(0.0022)	3.6737	(0.0021)	3.5215	(0.0021)
R^2	0.9604		0.8960		0.2559	
		Based on Order-Dependent Estimates: *Persons First*				
Firm-Size Category Average α	1.1372	(0.0045)	1.3224	(0.0042)		
Firm-Size Category Average ψ	0.9217	(0.0085)			1.7395	(0.0079)
Intercept	3.6370	(0.0022)	3.6674	(0.0021)	3.3665	(0.0019)
R^2	0.9990		0.8950		0.4327	

Notes: The dependent variable is the 25 firm-size category effects estimated by least squares controlling for labor force experience (through quartic), seniority, region, year, education (eight categories) and sex (fully interacted). See Table III for full results. The independent variables are the firm-size category averages for the indicated firm-specific compensation policy, adjusted for the same independent variables. The time period is 1976–1987.

TABLE IX

SUMMARY STATISTICS FOR FIRMS
ANNUAL AVERAGES OVER ALL YEARS FOR WHICH THE FIRM DOES BUSINESS 1978–1988
(weighted by inverse sampling probability)

Variable Definition	Mean	Std Dev
Order-Independent Estimates		
Average Predicted Effect of x Variables ($x\beta$) at the Firm	0.3569	0.2586
Average Individual Effect, Unobserved Factors (α) at the Firm	−0.0575	0.6626
Average Education Effect ($u\eta$) of Employees at the Firm	3.8889	0.2757
ϕ, Firm Effect Intercept	−0.1791	1.0279
γ, Firm Effect Seniority Slope	0.0156	0.1167
Order-Dependent Estimates: Persons First		
Average Predicted Effect of x Variables ($x\beta$) at the Firm	0.3906	0.2420
Average Individual Effect, Unobserved Factors (α) at the Firm	−0.0549	0.6446
Average Education Effect ($u\eta$) of Employees at the Firm	3.8503	0.2836
ϕ, Firm Effect Intercept	−0.0196	0.2707
γ, Firm Effect Seniority Slope	0.0027	0.0775
γ_2, Firm Effect Change in Slope at 10 Years	−0.0031	0.1728
Other Firm Characteristics		
Number of Employees Sampled at Firm	34.2950	610.4800
Employment at December 31st (thousands)	0.1097	1.6789
Real Total Assets (millions FF 1980)	59.4769	3,938.9800
Operating Income/Total Assets	0.1254	0.4544
Value Added/Total Assets	1.0051	1.8889
Real Total Compensation (millions FF 1980)	1.3260	2.3570
Real Value Added/Employee (thou. FF 1980)	106.7672	936.5212
Real Total Assets/Employee (thou. FF 1980)	363.0707	21,067.5500
(Engineers, Professionals and Managers)/Employee	0.2362	0.4072
Skilled Workers/Employee	0.5414	0.5255
Log(Real Total Assets)	1.7711	3.3558
Log(Real Value Added/Employee)	4.5215	1.1050
Log(Real Sales/Employee)	5.5673	2.0139
Log(Total Employment at December 31)	−3.0262	2.1109
Log(Real Total Assets/Employee)	4.7972	2.2710
Age of Firm ($N = 7{,}385$)	19.5023	23.0331
Number of Firms	14,717	

Notes: Order-independent estimates are based upon $x\beta$, α, $u\eta$ estimated with persons first (conditional on Z) and ϕ and γ estimated with firms first (conditional on Z). Order dependent estimates are all based upon persons first ($x\beta$, α, and $u\eta$) and firms (ϕ, γ, and γ_2) second.

ordinary least squares analysis in that much less of the variance is explained (between 0.30 and 0.55). To assess the quality of the different methods in estimating the β coefficients of the time-varying observable personal characteristics, we used Hausman (1978) tests to compare the coefficients obtained from the different methods with those obtained using the consistent method. Once again, all methods that include a person effect—the conditional method with

TABLE X

Generalized Least Squares Estimates of the Relation Between Productivity, Profitability and Compensation Policies

Dependent Variable:	Log (VAdded/Worker)		Log(Sales/Employee)		Operating Inc./Capital	
Independent Variable	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Based on Order-Independent Estimates						
Average Predicted Effect of x Variables ($x\beta$)	0.4937	(0.0270)	0.3050	(0.0393)	0.0670	(0.0151)
Average Individual Effect (α)	0.2234	(0.0108)	0.0809	(0.0156)	0.0081	(0.0060)
Average Education Effect ($u\eta$)	0.1338	(0.0254)	−0.0057	(0.0369)	−0.0107	(0.0143)
ϕ, Firm Effect Intercept	0.0371	(0.0084)	0.0054	(0.0122)	0.0138	(0.0047)
γ, Firm Effect Seniority Slope	−0.1210	(0.0582)	−0.1751	(0.0847)	−0.0028	(0.0328)
(Engineers, Professionals, Managers)/Employee	0.3428	(0.0238)	0.1773	(0.0346)	−0.1303	(0.0126)
(Skilled Workers)/Employee	0.1226	(0.0177)	0.3065	(0.0257)	0.0061	(0.0099)
Log(Capital/Employee)	0.2470	(0.0037)	0.5536	(0.0054)		
Intercept	3.0206	(0.1055)	0.1065	(0.1533)	0.1897	(0.0579)
Based on Order-Dependent Estimates: Persons First						
Average Predicted Effect of Variables ($x\beta$)	0.6057	(0.0310)	0.4833	(0.0494)	0.0569	(0.0161)
Average Individual Effect (α)	0.2617	(0.0118)	0.1623	(0.0188)	0.0102	(0.0061)
Average Education Effect ($u\eta$)	0.0725	(0.0275)	−0.0674	(0.0437)	−0.0036	(0.0143)
ϕ, Firm Effect Intercept	0.1240	(0.0343)	0.1128	(0.0546)	0.0415	(0.0179)
γ, Firm Effect Seniority Slope	0.1492	(0.1195)	0.2852	(0.1902)	0.0571	(0.0623)
γ_2, Firm Effect Change in Slope at 10 Years	−0.0485	(0.0428)	−0.1107	(0.0681)	−0.0264	(0.0223)
(Engineers, Tech., Managers)/Employee	0.6815	(0.0247)	0.8989	(0.0394)	−0.1267	(0.0126)
(Skilled Workers)/Employee	0.2167	(0.0190)	0.4979	(0.0302)	0.0094	(0.0099)
Log(Capital/Employee)	0.1017	(0.0025)	0.2290	(0.0039)		
Intercept	4.3985	(0.1126)	2.9784	(0.1791)	0.1664	(0.0586)

Note: Models were estimated using 14,717 firms with complete data. All regressions include a set of 2-digit industry effects.

persons first or order independent, the within-persons method, and the within-persons with limited firm effects method—perform better than methods without person effects.[35] Both ordinary least squares and within-firm estimates yield much higher χ^2 statistics, indicating that our preferred model must include person effects that are not orthogonal to the time varying effects in the model, including the conditioning effects $Z\lambda$.

[35] All the χ^2 statistics in models with person effects are around 3,500. In all cases, the statistic has 28 degrees of freedom. Hence, none of these models pass the test according to classical criteria. The models are also rejected using a simple Bayes-Schwartz criterion. However, given the large number of observations, we are likely to reject any model using these criteria. Hence, we use this test statistic as a measure of proximity of the β estimates to the consistent ones.

TABLE XI

GENERALIZED LEAST SQUARES ESTIMATES OF THE RELATION BETWEEN FACTORS USE AND COMPENSATION POLICIES

Independent Variable	Dependent Variable: Log(Employees)	Log(Real Capital)	Log(Capital /Employee)	EPM /Employee	Skilled W /Employee	Unskilled W /Employee
Based on Order-Independent Estimates						
Average Predicted	0.2586	1.0369	0.7783	0.1420	0.0542	−0.1962
Effect of x $(x\beta)$	(0.0675)	(0.0971)	(0.0600)	(0.0110)	(0.0140)	(0.0132)
Average Individual	0.2967	0.7673	0.4705	0.1197	−0.0284	−0.0913
Effect (α)	(0.0267)	(0.0384)	(0.0237)	(0.0043)	(0.0054)	(0.0051)
Average Education	0.4380	0.5479	0.1100	0.2974	−0.1060	−0.1915
Effect $(u\eta)$	(0.0638)	(0.0918)	(0.0567)	(0.0102)	(0.0130)	(0.0123)
ϕ, Firm Effect Intercept	−0.2654	−0.2898	−0.0244	0.0315	0.0152	−0.0468
	(0.0212)	(0.0304)	(0.0188)	(0.0035)	(0.0044)	(0.0042)
γ, Firm Effect Seniority	0.4305	0.4149	−0.0156	0.0909	−0.0147	−0.0762
Slope	(0.1465)	(0.2106)	(0.1300)	(0.0241)	(0.0306)	(0.0290)
(Eng., Prof., Managers)/	−0.0479	2.0645	2.1123			
Employee	(0.0565)	(0.0812)	(0.0501)			
(Skilled Workers)/	−0.2505	0.1075	0.3580			
Employee	(0.0444)	(0.0638)	(0.0394)			
Intercept	−3.6868	2.6123	6.2991	−0.7097	0.8567	0.8530
	(0.2587)	(0.3719)	(0.2296)	(0.0420)	(0.0534)	(0.0506)
Based on Order-Dependent Estimates: Persons First						
Average Predicted	0.2541	1.0205	0.7665	0.1142	0.0628	−0.1770
Effect of x $(x\beta)$	(0.0724)	(0.1036)	(0.0638)	(0.0117)	(0.0150)	(0.0142)
Average Individual	0.2764	0.7454	0.4690	0.1231	−0.0316	−0.0914
Effect (α)	(0.0273)	(0.0391)	(0.0241)	(0.0043)	(0.0055)	(0.0052)
Average Education	0.3478	0.4076	0.0598	0.3307	−0.0964	−0.2343
Effect $(u\eta)$	(0.0643)	(0.0921)	(0.0567)	(0.0101)	(0.0129)	(0.0122)
ϕ, Firm Effect Intercept	0.3748	0.7618	0.3869	0.0057	−0.0052	−0.0005
	(0.0802)	(0.1148)	(0.0707)	(0.0131)	(0.0167)	(0.0158)
γ, Firm Effect Seniority	−0.0262	0.5277	0.5539	0.0835	−0.0303	−0.0532
Slope	(0.2798)	(0.4005)	(0.2467)	(0.0456)	(0.0582)	(0.0553)
γ_2, Firm Effect Change	0.0011	0.0497	0.0486	−0.0314	0.0140	0.0174
in Slope at 10 Years	(0.1002)	(0.1435)	(0.0884)	(0.0164)	(0.0209)	(0.0198)
(Engi., Prof., Managers)/	−0.1181	2.0038	2.1219			
Employee	(0.0568)	(0.0812)	(0.0500)			
(Skilled Workers)/	−0.2947	0.0707	0.3654			
Employee	(0.0445)	(0.0637)	(0.0392)			
Intercept	−3.4129	3.0371	6.4499	−0.8485	0.8309	1.0176
	(0.2630)	(0.3765)	(0.2319)	(0.0423)	(0.0539)	(0.0512)

Notes: The models were estimated using the 14,717 firms with complete data. All equations include a set of 2-digit industry effects. Standard errors in parentheses.

TABLE XII

PROPORTIONAL HAZARDS ESTIMATES OF THE RELATION BETWEEN FIRM SURVIVAL AND COMPENSATION POLICIES

Independent Variable	Parameter Estimate	Standard Error	Risk Ratio
Based on Order-Independent Estimates			
Average Predicted Effect of x ($x\beta$)	2.2163	(0.5821)	9.1730
Average Individual Effect (α)	-0.5874	(0.2100)	0.5560
Average Education Effect ($u\eta$)	-2.3441	(0.5327)	0.0960
ϕ, Firm Effect Intercept	0.3833	(0.1579)	1.4670
γ, Firm Effect Seniority Slope	1.2239	(1.0215)	3.4000
(Eng., Prof., Managers)/Employee	0.2328	(0.3689)	1.2620
(Skilled Workers)/Employee	0.2065	(0.2917)	1.2290
Based on Order-Dependent Estimates: Persons First			
Average Predicted Effect of x ($x\beta$)	2.0751	(0.6241)	7.9650
Average Individual Effect (α)	-0.5327	(0.2064)	0.5870
Average Education Effect ($u\eta$)	-1.8615	(0.5398)	0.1550
ϕ, Firm Effect Intercept	-0.5909	(0.5356)	0.5540
γ, Firm Effect Seniority Slope	1.6497	(2.4598)	5.2050
γ_2, Firm Effect Change in Slope at 10 Years	0.3592	(0.6677)	1.4320
(Eng., Prof., Managers)/Employee	0.4096	(0.3699)	1.5060
(Skilled Workers)/Employee	0.3372	(0.2926)	1.4010

Notes: Negative coefficients indicate a reduced probability of firm death. This model was estimated using the 7,382 firms with known birth dates. The model includes a set of 2-digit industry effects.

We also computed the specification test shown in equation (3.40), which tests the hypothesis that the conditioning variables Z used to compute the column labelled "Conditional Method Persons First" are adequate to represent the covariance between personal characteristics, both measured and unmeasured, and firm effects. The computed statistic is 21,000 with 48 degrees of freedom. Since the conditioning variables have the most effect on the results when firm effects are estimated using the order-independent method, we conclude that this large χ^2 statistic, in conjunction with the component correlation analysis we discuss below, is evidence that the order-independent estimated firm effects are less reliable than the firm effects from the order-dependent "persons first" method. Of course, with the large sample sizes in this analysis, it is also the case that the large value of this statistic can be interpreted as having enough data to reject (unsurprisingly) a low-dimensional simplification of the covariance between X, D, and F. In spite of the data evidence that one should permit all of the effects to be correlated and that one should estimate person effects first in the conditional method, we present all of our results using both the order-independent method and the order dependent "persons first" methods. None of our conclusions are affected by our choice of conditional estimator.

5.2.2. *Male-female wage differentials*

Comparing the results for men and women in Table III, we note that there is less variation in the $\hat{\beta}$ across estimation methods for women than for men.[36] We also note that the gender gap is decreasing over our period of analysis according to the least squares estimates of the time effects with no person or firm effects. However, changes in the composition of the work force must have been an important determinant of this trend because, when person effects are included, the estimates of the time effects are virtually identical for the two sexes. Thus, given the overall difference between men and women in the French labor market, once we control for personal heterogeneity, there is no evidence of declining or increasing male-female wage differentials. As usual, the experience profile for women is flatter than for men, regardless of the method of estimation, even though, for our data source, the measure of labor force experience excludes within-sample periods of nonemployment.

5.3. *Discussion of the Estimated Person and Firm Effects*

Table IV contains descriptive statistics for the components of real compensation implied by the estimated parameters from both of our conditional method specifications, estimated separately for each sex. Table V contains pooled summary statistics and corrected correlations for all of the components of real compensation and for two different conditional estimation methods (order independent and order dependent "persons first"). The table also contains the different estimates of the seniority effects based on the estimation techniques presented in Table III. For both males and females, the standard deviations of the individual-effect, θ, and its components α and $u\eta$, are the same order of magnitude as the firm effects, ψ, for the order-independent method and substantially larger than those of the firm effects for the order-dependent method with persons first. As noted in Table III, the complete parameterization explains about 80% of the variation in real annualized earnings; thus, the person-specific component of variance is clearly important. The firm-specific component of variance is less important but still a major source of variation in the compensation data.

5.3.1. *Specification checks based on correlations among the heterogeneity components*

To further compare the different estimation methodologies, Table V shows the correlations among the components of person and firm heterogeneity as estimated using order-independent, order-dependent (both ways), within persons

[36] This statement is based upon the average variation in the coefficients for men versus those for women from the estimates in columns labelled "Consistent Method Person and Firm Effects," "Conditional Method Persons First," "Within Persons No Firm Effects," "Within Persons Limited Firm Effects," and "Within Firms No Person Effects."

with firm specific intercepts and seniority slopes for the 115 largest firms, and consistent methods. This table is particularly complicated and some care is required to read it properly. The correlation coefficients reported in the table are all computed to be representative of persons; hence, we use the full sample of 5,305,108 observations for all methods except for correlations with the full least squares estimates with limited firm effects, where the number of observations is equal to the 695,077 person-years for which the firm coefficients are available.

The panel labelled "Firm Effects" contains correlations of the components of the firm effects, ϕ and γ, by method of estimation. In the "Firm Effects" panel the order-independent estimates (based on equation (3.16)) are conditional on Z but exclude person effects; hence, they are equivalent to order-dependent "firms first" estimates, conditional on Z. In this same panel the order-dependent estimates are persons first, conditional on Z. The full least squares estimates using the 115 largest firms show the firm effects from the appropriate equation reported in Table III. Finally, in the "Firm Effects" panel, the consistent estimates are based on equation (3.8). Note that consistent estimates of the ϕ component of the firm effect are not available.

In the panel labelled "Person Effects," we report correlation coefficients based upon the order-independent estimates in equation (3.15), which are equivalent to order-dependent estimates with persons first. In this same panel we report person effect estimates from the order-dependent method with firms first.[37] Finally, the person effects from the model labelled "Full Least Squares Estimates Using the 115 Largest Firms," are based on the estimates reported in Table III. Note that consistent estimates of the person effects are not available.

5.3.2. *Specification checks based on firm effects, including heterogeneous seniority slopes*

Consider first the firm-specific intercept, ϕ. The results in Table IV show that, for both sexes, the standard deviation of the estimated firm effects is very large. The data evidence that the complete firm effect, ψ, is heterogeneous is particularly compelling when we combine the results shown in Table IV with the formal specification analyses we showed in Table III. Furthermore, the results in Table V show that order-dependent and order-independent methods give very different results for the firm effect since the correlation between the two estimates of ϕ is only 0.16. For the subsample of persons employed in the largest firms, the full least squares solution for γ appears to be closest to the consistent method (see below). Thus, we assess the quality of the estimates of ϕ obtained by the order-dependent and order-independent methods by comparing them with the full least squares solution for this subsample. Using this criterion, ϕ as esti-

[37] Because, as the reader will see shortly, these estimates perform very poorly, we do not report any other estimates based on the order-dependent method with firms first.

mated by the order-independent method is weakly correlated with ϕ as estimated with limited firm effects. On the other hand, the ϕ estimated by the order-dependent "persons first" method is strongly correlated with the ϕ estimated with limited firm effects (correlation of 0.67). Thus, the evidence based on ϕ favors the order-dependent "persons first" conditional estimation method. For clarity we stress that both conditional methods imply very substantial firm effects. The conclusion from this specification discussion is that the similarity between the full least squares estimates of ϕ (for the 115 largest firms) and the order-dependent "persons first" estimates indicates that the order-independent estimates of ϕ confound the pure firm-specific intercept with the average person effect within the firm.

Considering next the seniority coefficients, γ, all methods in which we allow these returns to vary across firms show that the standard deviation of the estimated seniority slopes is large, at least three times the mean. Our results, therefore, strongly suggest that earnings equations should have a firm effect with at least a firm-specific intercept and seniority slope. The various estimation methods, however, also show considerable variability in γ across techniques. The average seniority coefficient is about 0.01 whenever the estimation method excludes person effects (the order-independent method,[38] ordinary least squares, within firms, within industry and within size class). The average seniority coefficient decreases to near zero when person effects are included (order-dependent persons first, within persons, and 115 largest firms). The consistent method, which includes person effects, gives results closer to the models that exclude person effects—around 0.01 for the average seniority slope.

To continue our discussion of the seniority effects, consider the correlation among our estimates of this component of firm heterogeneity. Because we have consistent estimates of the seniority coefficients, it is useful to examine the correlation of γ from the consistent method with the other estimates. First, the correlation of the consistent γ with the γ estimated in the order-independent method is quite low (0.09). The correlation of the consistent γ with the one estimated by the order-dependent "persons first" method is only slightly larger. However, on the restricted sample of individuals for which the full least squares solution has been implemented, the correlation of the consistent γ with the full least squares γ is quite high (0.51). In fact, the γ estimated using the full least squares solution with the 115 largest firms is well-correlated with all of the methods of recovering γ over the subsample for which this estimate is available. Hence, for the largest firms, the seniority slope coefficients seem to be reasonably estimated by any of our methods. However, for the other, smaller, firms, no estimation method appears to dominate in a clear-cut fashion if one relies only on γ to assess the methodology.

[38] In the order-independent method the firm effects are estimated without first eliminating person effects; thus, they exclude person effects.

5.3.3. *Implications of heterogenous seniority slopes*

As we noted in Section 2, by considering the possibility of differential returns to seniority as a part of the firm effect, we can provide some direct evidence on the debate surrounding the interpretation of the average seniority effect. Using our consistent estimates of the return to seniority, γ, we find that the average return to a year of seniority is just over 0.01 for both sexes. This estimate is lower than Topel's (1991) result but consistent with Brown's (1989) results when he includes person effects. The heterogeneity in our consistent estimates suggests that some of the difference between our results and Topel's may be due to correlation between the heterogenous firm effect and the person effects. The fact that our results are closer to Brown's supports this conclusion because Brown's seniority effect is heterogenous—the magnitude of the return to seniority depends upon characteristics of the job—and he permits correlation between this heterogeneity and his person effect. Brown, on the other hand, does not allow for the possibility of firm-specific intercepts, except as reflected in the job characteristics he used to model the heterogeneity in the return to seniority. Although we cannot use our consistent technique to address this question, we note that, for all the preferred estimates of γ, there is a negative correlation between γ and the associated estimate of ϕ. This negative correlation indicates that the firm-specific intercept and the firm-specific seniority slope are negatively correlated, a result predicted by Becker and Stigler (1974) and Lazear (1979).

5.3.4. *Specification checks based on person effects*

Consider now the correlation between the different estimates of α. An argument similar to the analysis we used for ϕ shows that the α's estimated with persons first are better than those estimated with firms first. In the estimation of α, the order-independent estimates are mathematically identical to the order-dependent "persons first" estimates, conditional on Z. The alternative method is to consider the order-dependent "firms first" estimation of α. We note that the correlation between the order-independent estimates and the full least squares solution for the 115 largest firms is 0.99, while the order-dependent "firms first" estimates are only correlated 0.58 with the order-independent estimates and 0.60 with the full least squares solution for the 115 largest firms. These correlations indicate that the order-dependent: "firms first" estimates of α are not capturing the pure person effect as reliably as either of the other two alternatives shown in the "Person Effects" panel of Table V.

5.3.5. *Implications of the correlations among compensation components*

Table VI shows the intercorrelations of the different components of compensation, first for the order-independent method, then for the order-dependent "persons first" method. Both methods indicate that α, the unobservable part of the individual effect, is the component of compensation that is most highly correlated with log real annual total compensation (0.80 or 0.73 depending on the method).[39] The firm components are much less important in the determination of total compensation (0.21 or 0.26 depending on the method). Using the order-independent estimates, the α component of the person effect and the ϕ component of the firm effect are positively correlated 0.15. The estimated correlation is 0.08 using the order-dependent "persons first" estimates. In either case, the estimated correlation between firm and personal heterogeneity is not large. Also notice that, although the firm-specific intercept, ϕ, and the α-component of the person effect are positively correlated, the firm-specific intercept is negatively correlated with the seniority slope (-0.07 order independent and -0.56 order dependent "persons first"). In both methods, the correlation between observables and compensation appears to be smaller than the correlation between unobservables and compensation. The correlation between compensation and education, $u_i\eta$, is around 0.4 and the correlation between compensation and the time-varying individual characteristics, $x_{it}\beta$, is around 0.3, for both methods shown in the table. Furthermore, $x_{it}\beta$ is only weakly negatively correlated with the unobservable α.[40]

5.3.6. *Summary of the evidence from the estimation results on person and firm heterogeneity*

After reviewing the evidence of the quality of the different estimation methods, the following conclusions can be drawn. First, person effects tend to be more important than firm effects in explaining compensation variability. For the parameters α and β, the estimation methods with persons first are preferred by the data. However, there is no definitive evidence in favor of one estimation method over another for the firm effect ψ. On one hand, the order-dependent "persons first" method tends to give results (on ϕ) that are more highly correlated with the consistent estimates. On the other hand, the order-independent method produces estimates of γ that are less correlated with those obtained by the consistent methodology. Hence, in what follows, we will examine the classical problems of labor economics that were mentioned in the motivation section using person effects that have been estimated first (i.e. person effects from the order-independent and order-dependent "persons first" methods) and firm effects from these same methods, in all cases conditional on Z. This provides us with results that reflect the widest range of possibilities regarding the appropriate estimate of the firm effect.

[39] As noted in the discussion of statistical methods, at the level of the individual the least squares estimate of the person effect is unbiased but inconsistent. Thus, the variance of $\hat{\alpha}_i$ as directly calculated from the summary measures consists of two components $\text{var}[\alpha_i] + \text{var}[\hat{\alpha}_i - \alpha_i]$, and similarly for θ_i. The variances used to calculate all correlations with α_i and θ_i in Table VI have been corrected by subtracting an estimate of $\text{var}[\hat{\alpha}_i - \alpha_i]$, $\text{var}[\hat{\theta}_i - \theta_i]$, respectively.

[40] Recall that $\hat{\alpha}_i$ is orthogonal to $u_i\hat{\eta}$ by construction.

5.4. *Inter-Industry Wage Differentials*

In Table VII, we implement the equation (2.7) derived in Section 2, which allows us to decompose the industry effects, estimated as shown in Table III, into the component due to pure firm effects and the component due to person effects. Notice that the two right-hand side components must be adjusted for the observables as in equation (2.7). Table VII uses industry-level averages of the individual and firm-specific components of compensation to explain the industry effect found in our raw individual data (taken from the regression controlling for labor force experience, seniority, region, year, education, and sex reported in the column labelled "Within Industry No Person Effects" in Table III) in the spirit of Dickens and Katz (1987). Since the industry-average person and firm effects, also adjusted for the same set of factors as reported in the Table III regression, almost fully account for the industry effects in a statistical sense ($R^2 = 0.85$ using person and firm effects drawn from the order-independent method and $R^2 = 0.96$ using measures drawn from the order-dependent method with persons first), the interesting question concerns the relative importance of individual heterogeneity (the α-component of the person effect, in particular) and firm heterogeneity (the ψ-component) as components of the industry effects. For both estimation methods for the firm effects, the separate influence of person effects and firm effects in explaining the industry effects is shown. The separate analyses confirm the relative importance of person, as compared to firm, effects. The third through sixth columns of Table VII present separate industry-level regressions using, first, industry-average α alone (columns 3 and 4) and, then, using industry-average firm effects alone (columns 5 and 6). It is clear from the fact that industry-average α alone explains 84% (92% with the order-dependent estimates with persons first) of the inter-industry wage variation, whereas the industry-average ψ component explains only 7% (25% with the order-dependent estimates), that individual effects, as measured statistically by α, are more important than firm-components, as measured by ψ, for explaining French inter-industry wage differentials.[41]

Figures 1 and 2 show graphically the important difference in the strength of the relation between industry effects and industry-average person and firm effects. Figure 1 plots the industry effects from equation (2.7), the dependent variable in Table VII, against the industry-average person effects. The figure also shows the fitted regression line. Figure 2 plots the same industry effects against the industry-average firm effects, again showing the fitted regression line. Both figures are based on the order-independent estimates. The relation between the raw industry effects and the industry-average person effects is clearly much stronger than the one between raw industry effects and the industry-average firm effect. The graphical display for the order-dependent "persons first" estimates shows the same results.

[41] As shown in Table VI, these two components are not highly correlated, so little of the industry-average person effect is "explained by" the industry-average firm effect in a statistical sense.

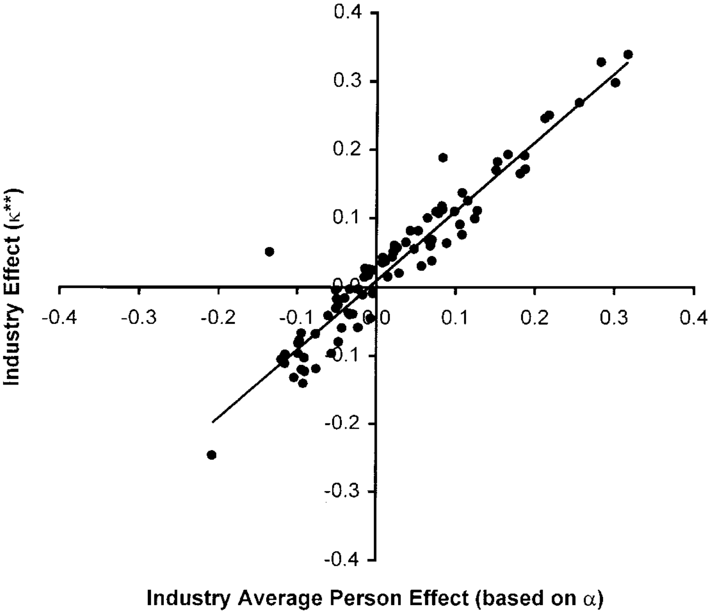

FIGURE 1.—Actual and predicted industry effects using industry-average person effects.

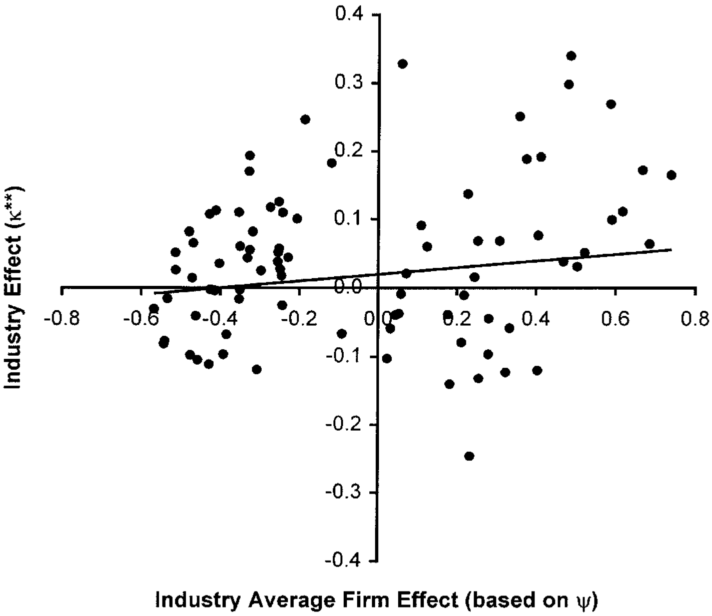

FIGURE 2.—Actual and predicted industry effects using industry-average firm effects.

5.5. *Firm-Size Wage Effects*

Table VIII presents a similar analysis for the firm-size effect based on equation (2.9). To implement this analysis, we constructed 25 firm-size categories. We then estimated the firm-size effects without controlling for person or firm effects, as in equation (2.9). Using calculations exactly parallel to those in Table VII, we constructed the appropriate weighted average person and firm effects within each firm-size category, conditional on the same X variables used in the other analyses. The complete set of X coefficients is shown in Table III in the column labelled "Within Firm Size No Person Effects." Table VIII shows that, for both methods of estimating the person and firm effects, the firm-size-average person effect is much better at explaining the firm-size wage effect than is the firm-size-average firm effect.[42] To more easily compare our results to others, Brown and Medoff (1989) in particular, we graph the raw firm-size wage effects against the log of firm size in Figure 3. The raw firm-size effects in our data strongly resemble the effects summarized by Brown and Medoff. The relation between firm size (log of employment at the firm) and compensation, controlling for the observable characteristics, follows a concave quadratic relation. Figure 3 also plots the average person effect (hollow boxes) within firm-size category. The average person effect can be seen to follow essentially the same quadratic function of log firm size and many of the average person effects are coincident with the solid dots representing the raw firm-size effect. Finally, Figure 3 shows the average firm effect (hollow triangles). The average firm effects do not follow the same concave quadratic function of log firm size as the other two effects. Indeed, the relation between the firm-size average firm effect and log firm size is slightly convex, with the largest positive average firm effect occurring in the largest firm-size category and the largest negative average firm effect occurring in the second largest firm size category. The effects plotted clearly show that average person effects are much more closely related to the firm size effects than average firm effects. The results shown are based on the order-independent estimates but are essentially the same for the order dependent estimates-persons first.

From these two analyses, we conclude that person effects are much more important in explaining inter-industry wage differentials and firm-size wage effects.

5.6. *The Economics of Human Resource Management*

We turn now to the analysis of the impact of the compensation structure on firm outcomes. To conduct this analysis we first computed the firm average of the different components of the compensation package as measured by our order-independent and order-dependent "persons first" methods. Hence, we

[42] As in the analysis of Table VII, the size effects used for the analysis in Table VIII come from the column in Table III labelled "Within Firm Size No Person Effects" and the size-class average person and firm effects have been adjusted for the same effects as found in the Table III regression.

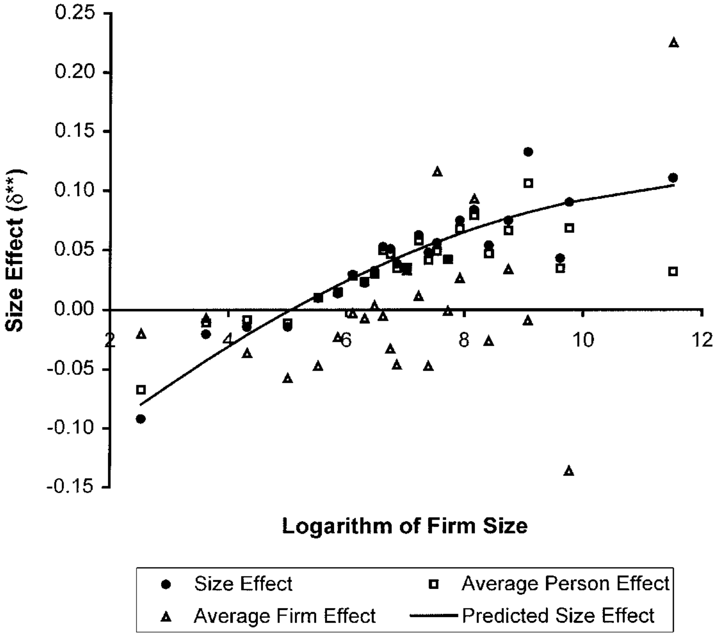

FIGURE 3.—Firm size effects related to firm-size average person and firm effects.

computed the average for each firm j of the part of compensation due to education $(u_i\eta)$, to time-varying observables $(x_{it}\beta)$, and to non-time-varying unobservables (α_i), using all observations (i,t) for which individual i was working in firm j at date t. The detailed formulas for this computation are described in the model section and the variables available for study are described in the data description.

Table IX presents summary statistics for the sample of firms (weighted to be representative of private industrial firms). Table X presents regression models of the logarithm of real value added per employee, real sales per employee (measures of productivity), and operating income as a proportion of total assets (a measure of performance). Results are reported for the order-independent and the order-dependent "persons first" methods. Using the firm-level compensation policy measures generated by our methods, we note that a larger value of the average component of the wage associated with time-varying characteristics $(x\beta)$ is associated with higher value-added and sales per worker and higher profitability for both estimation methods. A larger firm-average individual effect (α) is associated with a substantially larger value-added per employee and sales per employee but not with higher profitability. Once more, these results are consistent across estimation techniques. The part of the individual-effect related to education $(u\eta)$ is associated with higher value-added per worker but is not significant in the other two columns, irrespective of the estimation method. Higher firm-specific wages (ϕ) are associated with higher productivity (value-added per worker and sales per worker, albeit not with the order-independent method for this last variable) and with higher profitability.

The differences between the results based on the order-independent and order-dependent estimation methods, as shown in Table X, are most striking when looking at the impact of the seniority slope coefficient (γ). Using the order-dependent estimates, neither seniority slope is associated with higher (or lower) productivity or profitability. However, using the estimates from the order-independent method, it appears that there exists a negative association with firm productivity—firms that reward seniority the most tend to be the least productive.

The results in Table X can also be used to discuss the relation between firm level compensation policies and measurable outcomes in the context of hiring, rent-splitting, and efficiency wage models. Individuals with high opportunity wages, as captured by α, tend to work in firms with higher productivity per worker, as measured by either value-added per worker or sales per employee. Recall that the α-component of personal heterogeneity has been estimated using compensation as the dependent variable. Thus, it represents the market's valuation of this personal heterogeneity. It is thus not surprising that there is no profitability effect associated with α; however, for the same reason, the presence of an association between the observable characteristic component ($x\beta$) of compensation and profitability is puzzling, especially since the education component of individual heterogeneity has no measured association with profitability. The firm-specific effect in compensation, as measured by the firm-specific intercept ϕ, is associated with both higher productivity (value-added for either method and sales for the order-dependent measure) and higher profitability. This result can be interpreted as evidence consistent with some efficiency wage or rent-splitting activity in the labor market.

Table XI presents the results for the relations among our compensation measures and a variety of firm-level factor utilization rates. Results are also reported for both conditional estimation techniques. Larger values of the firm-average, time-varying component of compensation, $x\beta$, are associated with higher employment, capital, capital-labor ratio, proportion professional employment, and proportion skilled employment and with lower unskilled employment. The unobservable component of the individual effect, α, is positively associated with employment, capital, the capital-labor ratio, and the proportion of engineers, technical workers, and managers in the work force; and is negatively related to the shares of both skilled and unskilled workers. Larger values of the average education effect (the observable component of the individual effect, $u\eta$) are associated with higher employment, capital, and proportion professionals but with lower values of the proportion skilled. All of these results hold regardless of the estimation method.

The estimation method for the compensation components matters when examining the impact of the firm effects on these outcomes. Based on the order-dependent "persons first" method, the firm-specific intercept, ϕ, is strongly positively associated with employment, capital, and the capital-labor ratio; but is not associated with any components of the skill structure of the work force. A high firm-specific seniority slope is positively associated with the capital-labor ratio and slightly positively associated with the proportion of professional employees. Based on the order-independent method, all of the associations with ϕ that were positive using the order-dependent method are now negative (significantly for employment and capital, marginally for the capital-labor ratio); but the firm-specific seniority slope, γ, plays the role that the firm-specific intercept, ϕ, played with the other estimation method—it is positively associated with employment and capital. Furthermore, managerial and skilled employment are both positively associated with firm-specific effects. It appears that our two estimation techniques both capture similar effects, but their allocation to the fixed part and to the seniority part of firm-specific heterogeneity differ. This is confirmed by a look at Table V in which we see that ϕ from the order-dependent method is highly negatively correlated with γ from the order-independent method and that the ϕ from the order-independent method is somewhat negatively correlated with γ from the order-dependent method.

Finally, Table XII presents a proportional hazards analysis of the relation between the survival of firms and our estimated compensation components at the firm level.[43] Both components of the individual effect, α and $u\eta$, are associated with an increase in survival probability in a statistically significant manner. The effects related to firm-specific compensation factors are large but very imprecise, even though a high ϕ tends to decrease survival when using order-independent estimates. The effect associated with the firm average of observable personal characteristics, $x\beta$, is also associated with a decreased survival probability. The results are interesting when combined with those found in Table X. High ϕ is related to high profitability with both estimation methods, but is linked to lower probabilities of firm survival. On the other hand, high α is related to increased survival probabilities, but has no significant relation to profitability.

6. CONCLUSIONS

In Section 2 we identified six broad areas of labor economics that could be advanced by the study of matched longitudinal employer-employee data:

- the role of individual and firm heterogeneity in the determination of wage rates;
- the sources of inter-industry wage differentials;
- the sources of firm-size wage effects;
- the role of seniority, and heterogeneous returns to seniority in determining wage rates;
- the measurement of internal and external wage rates;
- the study of the economics of human resource management policies.

[43] We estimate the Cox proportional hazards model using as independent variables the non-time-varying measures shown in Table XII. The nonparametric baseline hazard was not estimated.

We believe that our analysis of the French compensation data, linked to the economic performance data of the employing firms has, indeed, shed considerable new light on these questions. To summarize, we found:

- Personal heterogeneity and firm heterogeneity were both important determinants of compensation, although personal heterogeneity appears to be substantially more important in these French data.
- Across 84 industries, the industry-average person effect, adjusted for inter-industry differences in observable characteristics, is much more important than the industry-average firm effect, similarly adjusted, for explaining the inter-industry wage differential.
- Across 25 employment-size categories, the firm-size wage effect in France is increasing at a decreasing rate and this effect is more closely predicted by a similar pattern in the firm-size-average person effect than by the firm-size-average firm effect, which does not mirror the raw firm-size effects at all.
- There is considerable evidence for heterogenous returns to seniority but the method of estimating the return to seniority affected the conclusion regarding the average return to one year of additional seniority. Returns to seniority are negatively correlated with firm-specific intercepts in the compensation relation.
- If we associate the person effect with an individual's external wage rate and the firm effect with that person's internal wage rate, there is very little correlation between these two measures, suggesting that models that focus on explanations for the individual heterogeneity (human capital) and models that focus on explanations for the firm heterogeneity (compensation design, incentives, bargaining) are addressing features of the labor market that do not have large interactions.
- Firms that hire "high-wage workers," those with above average person effects, are observed to have more productive work forces but no higher profitability. "High-wage firms," those that pay above average firm effects, are observed to have both more productive work forces and higher profits.

Of course, our analysis of the separate effects of individual and firm heterogeneity on wage rates and on firm compensation policies has also raised many new questions:

- Do the results for France generalize to other labor markets?
- If person effects are much more important than firm effects in explaining variation in compensation, do these same effects also explain employment mobility?
- If pure firm effects are not very important in the explanation of inter-industry wage differences, then why do other analyses that control for personal heterogeneity but not for firm heterogeneity appear to suggest otherwise?
- Does the observed relation between hiring "high-wage" workers and having higher productivity per worker mean that employer's hiring and selection methods should be studied more closely?
- Is the observed relation between being a "high-wage" firm and being both more profitable and more productive per worker evidence that efficiency wage models play a role in explaining inter-firm differences in compensation policies?

Although we have provided considerable new evidence on these outstanding questions, we believe that our results also provide the statistical basis upon which to begin the process of testing the relevance of agency, efficiency wage, search/matching, rent sharing and endogeneous mobility models as potential explanations for compensation outcome heterogeneity.

Dept. of Labor Economics, 259 Ives Hall, Cornell University, Ithaca, New York 14853-3901, U.S.A.; John.Abowd@cornell.edu; http://old-instruct1.cit.cornell.edu:8000/abowd-john/; CREST, Malakoff Cedex, France, and NBER, Cambridge, MA, U.S.A.,

CREST-INSEE, Dept. de la Recherche, 15, Bd. Gabriel Péri, 92245 Malakoff Cedex, France; kramarz@ensae.fr; and CEPR, London, U.K.,

and

CNRS, LAMIA-TEAM, Université de Paris 1 (Panthéon-Sorbonne), Maison des Sciences Economiques, 106-112, Bd. de l'Hôpital, 75647 Paris Cedex 13, France; margolis@univ-paris1.fr; and CREST, Malakoff Cedex, France

Manuscript received February, 1994; final revision received January, 1998.

STATISTICAL APPENDIX

In this appendix we state and prove the basic statistical results relating our estimation techniques and our analysis of the aggregation and suppression of effects to the standard least squares analysis of individual and firm effects that have been estimated in other contexts. The model is stated in equation (2.2) and the definitions that follow. We use the same notation in this appendix.

There are a total of J firms indexed by $j = 1, \ldots, J$. The function $\mathbf{J}(i,t)$ gives the identity of the employer for individual i in period t. For each individual i and each year $t = n_{i1}, \ldots, n_{iT_i}$, a row of the matrix F_0 contains an indicator variable for which the jth column contains the value 1 and all other columns contain the value 0, where $j = \mathbf{J}(i,t)$. The matrix F_0 is, thus, $N^* \times J$ and the associated vector of firm effects, ϕ, is $J \times 1$. A row of the matrix F_1 contains, for each individual i and each year $t = n_{i1}, \ldots, n_{iT_i}$, in the jth column the value of the individual's seniority in the firm $j = \mathbf{J}(i,t)$, s_{it}, and 0 in all other columns. A row of the matrix F_2 contains, for each individual i and each year $t = n_{i1}, \ldots, n_{iT_i}$, in the jth column the value of the individual's seniority in the firm $j = \mathbf{J}(i,t)$ less 10 if this value is positive and 0 otherwise, $\mathbf{T}_1(s_{it} - 10)$, and 0 in all other columns. The complete firm effect can thus be represented as

$$F\psi = F_0\phi + F_1\gamma + F_2\gamma_2$$

where $F \equiv [F_0 \;\; F_1 \;\; F_2]$ and $\psi \equiv [\phi' \;\; \gamma' \;\; \gamma_2']'$.

The error vector, ε, is $N^* \times 1$ and has the following properties:

$$\mathrm{E}[\varepsilon | X, D, F] = 0,$$

$$\mathrm{var}[\varepsilon | X, D, F] = \sigma^2 I_{N^*}.$$

Hence, the full regression equation for the model in the main text of the paper is given by

$$y = X\beta + D\theta + F_0\phi + F_1\gamma_1 + F_2\gamma_2 + \varepsilon. \tag{7.1}$$

For completeness we note that the N individuals constitute a simple random sample of the population of persons ever employed (outside the government sector) between the years 1976 and 1987 (except for 1981 and 1983, for which the data were not made available in a computerized sample). In general, the individuals were sampled if their birth dates fell in October of an even year.

Once sampled, an individual's complete private-sector employment history between the years 1976 and 1987 is available, again except for the years 1981 and 1983.

At most $P+N+(3J-1)$ effects in the full model are identified. The least squares estimator of the complete set of effects is given by

$$\begin{bmatrix}\hat{\beta}\\ \hat{\theta}\\ \hat{\psi}\end{bmatrix}=\begin{bmatrix}X'X & X'D & X'F\\ D'X & D'D & D'F\\ F'X & F'D & F'F\end{bmatrix}^{-}\begin{bmatrix}X'y\\ D'y\\ F'y\end{bmatrix}\tag{7.2}$$

where the notation $[\ ]^{-}$ represents any generalized inverse. The standard method of calculating the least squares estimates of the effects is to take deviations from the within-person means of the variables. This operation is accomplished by premultiplying both sides of equation (7.1) by the matrix $M_D\equiv[I-D(D'D)^{-1}\ D']$. The least squares estimator of the identifiable effects can, then, be restated as

$$\begin{bmatrix}\hat{\beta}\\ \hat{\psi}\end{bmatrix}=\begin{bmatrix}X'M_DX & X'M_DF\\ F'M_DX & F'M_DF\end{bmatrix}^{-}\begin{bmatrix}X'M_Dy\\ F'M_Dy\end{bmatrix}.\tag{7.3}$$

It is because the off-diagonal submatrix $X'M_DF$ is neither null, patterned, nor sparse that we cannot directly compute the solution (7.3). Furthermore, even if we use the consistent estimators $\tilde{\beta}$ and $\tilde{\gamma}$ from equations (3.9) and (3.10) and set $\gamma_2=0$, because of the presence of the person effects, the consistent estimator for $\tilde{\phi}$ based upon equation (7.3) is

$$\tilde{\phi}=(F_0'M_DF_0)^{-}F_0'M_D(y-X\tilde{\beta}-F_1\tilde{\gamma}),$$

which still requires the solution of a system of J equations that is neither diagonal, patterned, nor sparse.

Computation of the Conditional Estimates

The calculation of the solution to equations (3.17) and (3.18) for the order-independent and order-dependent (persons first) methods do not present any problems. The calculation of equation (3.21) for the order-independent and order-dependent "firms first" methods is, however, more complicated. For the order dependent method with firms first we leave $F_2\gamma_2$ in the model; however, we do not attempt the order independent calculations with $F_2\gamma_2$ in the model.

To calculate $\hat{\pi}$, we reorganize the columns of F so that the columns of F_0 and F_1 from the same firm are adjacent. Next, we sort the matrix F so that the observations are grouped by firm from $j=1,\ldots,J$. Denote the reorganized F matrix by F^* and denote the conformably reorganized y and Z matrices by y^* and Z^*, respectively. Note that the cross-product matrix $F^{*\prime}F^*$ is block diagonal with J blocks, each one 2×2, and a typical block is given by

$$\begin{bmatrix}N_j & \sum_{\mathbf{J}(i,t)=j} s_{it}\\ \sum_{\mathbf{J}(i,t)=j} s_{it} & \sum_{\mathbf{J}(i,t)=j} s_{it}^2\end{bmatrix}$$

where

$$N_j\equiv\sum_{\forall(i,t)}1[\mathbf{J}(i,t)=j],$$

the notation $\Sigma_{\mathbf{J}(i,t)=j}$ means to sum over all (i,t) such that $\mathbf{J}(i,t)=j$, and the function $1[A]=1$ if A is true and 0, otherwise. Similarly, the cross-product matrix $F^{*\prime}Z^*$, which is $2J\times Q$, has the structure

$$\begin{bmatrix} \sum\limits_{\mathbf{J}(i,t)=j} z_{(i,t)1} & \cdots & \sum\limits_{\mathbf{J}(i,t)=j} z_{(i,j)Q} \\ \sum\limits_{\mathbf{J}(i,t)=j} s_{it} z_{(i,t)1} & \cdots & \sum\limits_{\mathbf{J}(i,t)=j} s_{it} z_{(i,t)Q} \\ \cdots & \cdots & \cdots \\ \sum\limits_{\mathbf{J}(i,t)=j} z_{(i,t)1} & \cdots & \sum\limits_{\mathbf{J}(i,t)=j} z_{(i,t)Q} \\ \sum\limits_{\mathbf{J}(i,t)=j} s_{it} z_{(i,t)1} & \cdots & \sum\limits_{\mathbf{J}(i,t)=j} s_{it} z_{(i,t)Q} \end{bmatrix}.$$

The product $(F^{*\prime}F^{*})^{-1}F^{*\prime}Z^{*}$ is, therefore, a $2J \times Q$ matrix of firm-specific regression coefficients. A similar argument can be made for the coefficients associated with the regression of y^* on Z^*. Hence, the adjustment of Z^* with respect to F^* can be accomplished performing firm by firm regression of the appropriate rows of each column of Z^* on the appropriate columns of F^* and retaining the residuals to cumulate in the cross-product matrices $(Z^{*\prime}M_{F^*}Z^*)$ and $(Z^{*\prime}M_{F^*}y^*)$. Thus

$$\hat{\pi} = (Z^{*\prime}M_{F^*}Z^*)^{-1}(Z^{*\prime}M_{F^*}y^*),$$

where we note that it is not necessary to adjust y^* with respect to F^* as long as each column of Z^* has been adjusted (i.e. the matrix M_{F^*} is idempotent).

A computationally identical approach to the estimation of π may be obtained by directly solving the least squares equations associated with the solution of (3.20). To begin, notice that the least squares solution to this equation has the property

$$\sum_{\mathbf{J}(i,t)=j} [y_{it} - \phi_j - \gamma_j s_{it} - z_{it}\pi] = 0$$

for $j = 1, \ldots, J$, where the variable z_{it} is a row of the matrix Z. These J conditions imply

$$\hat{\phi}_j = \bar{y}_j - \hat{\gamma}_j \bar{s}_j - \bar{z}_j \hat{\pi} \tag{7.4}$$

where the notation

$$\bar{a}_j \equiv \frac{\sum_{\mathbf{J}(i,t)=j} a_{it}}{\sum_{\forall(i,t)} 1[\mathbf{J}(i,t)=j]}.$$

Next, consider the J orthogonality conditions associated with the variable s_{it}, which imply

$$\sum_{\mathbf{J}(i,t)=j} \left(y_{it}s_{it} - \hat{\phi}_j s_{it} - \hat{\gamma}_j s_{it}^2 - z_{it}s_{it}\hat{\pi} \right) = 0$$

for $j = 1, \ldots, J$. Hence,

$$\hat{\gamma}_j = \frac{\sum_{\mathbf{J}(i,t)=j}[(y_{it} - \bar{y}_j)s_{it} - (z_{it} - \bar{z}_j)s_{it}\hat{\pi}]}{\sum_{\mathbf{J}(i,t)=j}(s_{it} - \bar{s}_j)s_{it}}. \tag{7.5}$$

Substituting equations (7.4) and (7.5) into (3.19) yields

$$\begin{aligned} & y_{it} - \bar{y}_j - \frac{[\sum_{\mathbf{J}(i,t)=j}(y_{it} - \bar{y}_j)](s_{it} - \bar{s}_j)1(\mathbf{J}(i,t)=j)}{\sum_{\mathbf{J}(i,t)=j}(s_{it} - \bar{s}_j)s_{it}} \\ & \quad = \left(z_{it} - \bar{z}_j - \frac{[\sum_{\mathbf{J}(i,t)=j}(z_{it} - \bar{z}_j)s_{it}]((s_{it} - \bar{s}_j)1(\mathbf{J}(i,t)=j))}{\sum_{\mathbf{J}(i,t)=j}(s_{it} - \bar{s}_j)s_{it}} \right)\pi. \end{aligned} \tag{7.6}$$

Thus, the least squares estimator of the $Q \times 1$ vector π is given by

$$\hat{\pi} = (\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'\tilde{y}$$

where the $1 \times Q$ vector

$$\tilde{z}_{it} \equiv z_{it} - \bar{z}_j - \frac{[\Sigma_{\mathbf{J}(i,t)=j}(z_{it} - \bar{z}_j)s_{it}]((s_{it} - \bar{s}_j)1(\mathbf{J}(i,t)=j))}{\Sigma_{\mathbf{J}(i,t)=j}(s_{it} - \bar{s}_j)s_{it}}$$

and the scalar

$$\tilde{y}_{it} \equiv y_{it} - \bar{y}_j - \frac{[\Sigma_{\mathbf{J}(i,t)=j}(y_{it} - \bar{y}_j)s_{it}](s_{it} - \bar{s}_j)1(\mathbf{J}(i,t)=j)}{\Sigma_{\mathbf{J}(i,t)=j}(s_{it} - \bar{s}_j)s_{it}}.$$

Finally, using either computational formula for the estimator $\hat{\pi}$, we have

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\gamma}_1 \\ \cdots \\ \hat{\phi}_J \\ \hat{\gamma}_J \end{bmatrix} = (F^{*\prime}F^*)^{-1}F^{*\prime}(y^* - Z^*\hat{\pi}), \tag{7.7}$$

which, once again, can be computed firm-by-firm using the appropriate columns of F^* and the appropriate rows of $(y^* - Z^*\hat{\pi})$, so that

$$\begin{bmatrix} \hat{\phi}_j \\ \hat{\gamma}_j \end{bmatrix} = \begin{bmatrix} N_j & \sum\limits_{\mathbf{J}(i,t)=j} s_{(i,t)} \\ \sum\limits_{\mathbf{J}(i,t)=j} s_{(i,t)} & \sum\limits_{\mathbf{J}(i,t)=j} s_{(i,t)}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum\limits_{\mathbf{J}(i,t)=j} \left(y_{(i,t)}^* - \sum\limits_{q=1}^{Q} z_{(i,t)q}^* \hat{\pi}_q \right) \\ \sum\limits_{\mathbf{J}(i,t)=j} \left(y_{(i,t)}^* - \sum\limits_{q=1}^{Q} z_{(i,t)q}^* \hat{\pi}_q \right) s_{(i,t)} \end{bmatrix}.$$

Least Squares and Our Conditional Methods

Consider next the relation between our conditional method estimators and the conventional least squares estimator. Because this appendix contains the proofs of the claims in the paper, we use the full model in equation (7.1). The conditional method matrix Z can be expressed as

$$\left[\begin{array}{c|c|c} \bar{x}_1 f_{\mathbf{J}(1,n_{11})} & \bar{x}_1 s_{1n_{11}} f_{\mathbf{J}(1,n_{11})} & \bar{x}_1 \mathbf{T}_1(s_{1n_{11}} - 10) f_{\mathbf{J}(1,n_{11})} \\ \cdots & \cdots & \cdots \\ \bar{x}_1 f_{\mathbf{J}(1,n_{1T_1})} & \bar{x}_1 s_{1n_{1T_1}} f_{\mathbf{J}(1,n_{1T_1})} & \bar{x}_1 \mathbf{T}_1(s_{1n_{1T_1}} - 10) f_{\mathbf{J}(1,n_{1T_1})} \\ \cdots & \cdots & \cdots \\ \bar{x}_N f_{\mathbf{J}(N,n_{N1})} & \bar{x}_N s_{Nn_{N1}} f_{\mathbf{J}(1,n_{N1})} & \bar{x}_N \mathbf{T}_1(s_{Nn_{N1}} - 10) f_{\mathbf{J}(1,n_{N1})} \\ \cdots & \cdots & \cdots \\ \bar{x}_N f_{\mathbf{J}(N,n_{NT_N})} & \bar{x}_N s_{Nn_{NT_N}} f_{\mathbf{J}(1,n_{NT_N})} & \bar{x}_N \mathbf{T}_1(s_{Nn_{NT_N}} - 10) f_{\mathbf{J}(1,n_{NT_N})} \end{array}\right]$$

where $\bar{x}_i$ are the rows of $(D'D)^{-1}D'XC$, C is a $P \times (Q/3)$ matrix that selects $Q/3$ columns of X to place in the Z matrix, $\mathbf{T}_1(z)$ is the first order spline basis function defined in the text of the article, and all other variables are defined above. Hence, Z is $N^* \times Q$.

We express the projection of $F\psi$ on Z as

$$F_0\phi + F_1\gamma_1 + F_2\gamma_2 = Z\lambda + \nu$$

where the vector λ is $Q\times 1$ and the error process ν is defined as the component of the firm effect that is orthogonal to $Z\lambda$. The statistical equation substituting the projection of the firm effects for the actual firm effects is given by

$$y = X\beta + D\theta + Z\gamma + \varepsilon + \nu$$

with an error process $\varepsilon+\nu$ with the following properties:

$$\mathrm{E}[\varepsilon+\nu|X,D,Z]=0,$$

$$\mathrm{var}[\varepsilon+\nu|X,D,Z]=\sigma_1^2 I_{N*},$$

where σ_1^2 is the variance of $\varepsilon_{it}+\nu_{it}$ for all (i,t) that are part of N^*. As a direct consequence of this statistical model we are assuming the following orthogonality condition:

$$\begin{bmatrix} X'(F_0\phi+F_1\gamma_1+F_2\gamma_2-Z\lambda) \\ D'(F_0\phi+F_1\gamma_1+F_2\gamma_2-Z\lambda) \end{bmatrix}=0, \tag{7.8}$$

which means that the columns of Z should be chosen to maximize the correlation of X and D with $F_0\phi+F_1\gamma_1+F_2\gamma_2$. We calculate the within-person least squares estimator for β and λ using the formula

$$\begin{bmatrix} \hat{\beta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} X'M_D X & X'M_D Z \\ Z'M_D X & Z'M_D Z \end{bmatrix}^{-} \begin{bmatrix} X'M_D y \\ Z'M_D y \end{bmatrix}.$$

The proof of the consistency of this estimator follows directly from the condition (7.8) so that the asymptotic distribution of $[\,\hat{\beta}' \;\; \hat{\lambda}'\,]'$ is given by the usual least squares formulas.

Aggregation of Effects

We consider next the consequences of various aggregations and substitutions on the least squares estimators of the various effects in the model 2.2. The algebra for all of the aggregations considered in Section 2 is identical so we will discuss only the generic case in this appendix. An aggregation of the firm effect can be defined as an orthogonal decomposition of the firm effect into a part related to the aggregation and a part that represents the residual from this aggregation. We consider the industry aggregation given by the matrix A and the parameters κ, defined in Section 2.

The model (2.2) can be restated as

$$y = X\beta + D\theta + FA\kappa + (I_{N*} - FA(A'F'FA)^{-}A'F')F\psi + \varepsilon. \tag{7.9}$$

If the firm effects are omitted from the model, then the statistical error becomes

$$\zeta \equiv (I_{N*} - FA(A'F'FA)^{-}A'F')F\psi + \varepsilon. \tag{7.10}$$

By construction, the design matrices FA and $(I_{N*}-FA(A'F'FA)^{-}A'F')F$ are orthogonal. However, neither design matrix is orthogonal to X or D. Thus, the least squares estimates of the pure class effects, κ, suffer from an excluded variable bias when they are estimated in the absence of firm effects. Specifically, the within-person least squares estimator of the effects β^* and κ^* from equation (7.9) with the error term defined by equation (7.10) is

$$\begin{bmatrix} \hat{\beta}^* \\ \hat{\kappa}^* \end{bmatrix} = \begin{bmatrix} X'M_D X & X'M_D FA \\ A'F'M_D X & A'F'M_D FA \end{bmatrix}^{-} \begin{bmatrix} X'M_D X\beta + X'M_D FA\kappa + X'M_D\zeta \\ A'F'M_D X\beta + A'F'M_D FA\kappa + A'F'M_D\zeta \end{bmatrix}.$$

By direct calculation of the partitioned G-inverse we have

$$\operatorname*{plim}_{N\to\infty} \hat{\kappa}^* = \kappa - Q^{-}(A'F'M_D X(X'M_D X)^{-}X'M_D)(I_{N*} - FA(A'F'FA)^{-}A'F')F\psi$$

where

$$Q \equiv (A'F'M_D FA - A'F'M_D X(X'M_D X)^{-} X'M_D FA).$$

By inspection we note that the source of the inconsistency in the within-persons least squares estimator of the class effects κ is the covariance between the observed characteristics, X, and the part of the firm effects that is not correlated with the industry effects, $(I_{N*} - FA(A'F'FA)^{-}A'F')F$, conditional on the person effects D.

For completeness we note that if the pure class effect, say κ^{***}, is defined to be representative of firms, and not of individuals, then

$$\kappa^{***} \equiv (A'A)^{-} A'\psi.$$

Using this definition of the pure class effect, there will be an additional term in the probability limit of $\hat{\kappa}^{***}$ that gives the aggregation bias associated with estimating this pure class effect using the firm design matrix F and a sampling plan that is representative of persons. To our knowledge, none of the articles cited in this paper that estimate industry or size effects from samples that are representative of the population of employed individuals use a definition of a class effect that is representative of the population of firms.

Firm Effects That Depend on Firm-level Data

Suppose next that the firm effect, ψ, depends upon a non-time-varying characteristic of the firm over the sample period. Let the $J \times 1$ vector f contain the characteristic of firm j, less the grand mean, in each row. The grand mean should be calculated over the population of employed individuals so that the average firm effect in the population of persons remains zero. Because the parameters of our firm effects are constant over time, we cannot nest a model of time-varying firm characteristics in equation (2.2). The pure firm effects can be decomposed into the part that is linearly related to f and a residual from this linear relation:

$$\psi = f\delta + v$$

where δ is a scalar parameter relating the firm's characteristic to its firm effect and the $J \times 1$ vector v gives the residual from this projection. By an argument completely analogous to the one we used for pure classification effects, it can be shown that the within-person least squares estimator of δ is also inconsistent because X and v are not orthogonal, conditional on D. Specifically,

$$\operatorname*{plim}_{N \to \infty} \hat{\delta} = \delta + \frac{1}{q} f'FM_D(I_{N*} - X(X'M_D X)^{-} X')M_D F(\psi - f\delta)$$

where $q = f'F'M_D Ff - f'F'M_D X(X'M_D X)^{-}X'M_D Ff$.

DATA APPENDIX

This Appendix contains details of the definitions of variables, missing data imputation, and statistical calculations not reported in the text.

Education and School-Leaving Age

Our initial DAS file did not contain education information. We used supplementary information available for a strict subsample of the DAS (called the EDP, Echantillon Démographique Permanent) to impute the level of education of all other individuals in the DAS as described in Section 4. The education responses were grouped into 8 degree-level categories as shown in Data Appendix Table B1. EDP sample statistics for the men are in Data Appendix Table B2, and those for the women are in Data Appendix Table B3. The estimated logit equations are in Data Appendix Table B4 for men and Data Appendix Table B5 for women.

DATA APPENDIX TABLE B1

CLASSIFICATION OF FRENCH DEGREES AND U.S. EQUIVALENTS

Category	Degree	U.S. Equivalent
1	Sans Aucun Diplôme	No Terminal Degree
2	CEP	Elementary School
	DFEO	
3	BEPC	Junior High School
	BE	
	BEPS	
4	BAC (not F, G or H)	High School
	Brevet supérieur	
	CFES	
5	CAP	Vocational-Technical School (Basic)
	BEP	
	EFAA	
	BAA	
	BPA	
	FPA 1er	
6	BP	Vocational-Technical School (Advanced)
	BEA	
	BEC	
	BEH	
	BEI	
	BES	
	BATA	
	BAC F	
	BAC G	
	BAC H	
7	Santé	Technical College and
	BTS	Undergraduate University
	DUT	
	DEST	
	DEUL	
	DEUS	
	DEUG	
8	2ème cycle	Graduate School and Other
	3ème cycle	Post-Secondary Education
	Grande école	
	CAPES	
	CAPET	

Notes: Authors' adaptation of French degree codes appearing on the EDP (Echantillon démographique permanent).

DATA APPENDIX TABLE B2

EDP SAMPLE STATISTICS–MEN
(Std. Deviations in Parentheses)

Variable Name	Overall	Degree Category 1	2	3	4	5	6	7	8
$DOB_i < 1925$	0.188	0.254	0.295	0.160	0.136	0.055	0.098	0.063	0.186
	(0.391)	(0.435)	(0.456)	(0.367)	(0.343)	(0.228)	(0.297)	(0.243)	(0.389)
$1924 < DOB_i < 1930$	0.056	0.062	0.085	0.042	0.049	0.034	0.048	0.026	0.065
	(0.230)	(0.242)	(0.279)	(0.200)	(0.215)	(0.180)	(0.214)	(0.158)	(0.247)
$1929 < DOB_i < 1935$	0.097	0.109	0.120	0.067	0.068	0.081	0.095	0.054	0.101
	(0.296)	(0.311)	(0.325)	(0.250)	(0.252)	(0.273)	(0.293)	(0.226)	(0.301)
$1934 < DOB_i < 1940$	0.061	0.056	0.070	0.048	0.048	0.063	0.079	0.047	0.078
	(0.240)	(0.229)	(0.255)	(0.214)	(0.215)	(0.244)	(0.270)	(0.212)	(0.268)
$1939 < DOB_i < 1945$	0.094	0.070	0.091	0.075	0.098	0.117	0.133	0.118	0.149
	(0.292)	(0.256)	(0.287)	(0.264)	(0.298)	(0.322)	(0.340)	(0.323)	(0.356)
$1944 < DOB_i < 1950$	0.102	0.064	0.097	0.099	0.130	0.130	0.152	0.175	0.164
	(0.302)	(0.244)	(0.296)	(0.299)	(0.336)	(0.336)	(0.359)	(0.380)	(0.370)
$1949 < DOB_i < 1955$	0.159	0.095	0.132	0.166	0.245	0.224	0.217	0.288	0.201
	(0.365)	(0.293)	(0.339)	(0.372)	(0.430)	(0.417)	(0.412)	(0.453)	(0.401)
$1954 < DOB_i < 1960$	0.101	0.072	0.060	0.182	0.157	0.145	0.110	0.176	0.054
	(0.302)	(0.259)	(0.238)	(0.386)	(0.364)	(0.352)	(0.313)	(0.381)	(0.226)
$1959 < DOB_i < 1977$	0.141	0.218	0.050	0.160	0.069	0.151	0.068	0.052	0.003
	(0.348)	(0.413)	(0.218)	(0.367)	(0.253)	(0.358)	(0.251)	(0.224)	(0.056)
Works in	0.232	0.204	0.226	0.288	0.352	0.187	0.284	0.309	0.457
Ile de France	(0.422)	(0.403)	(0.418)	(0.453)	(0.478)	(0.390)	(0.451)	(0.462)	(0.498)
CSP62	0.263	0.357	0.282	0.188	0.157	0.199	0.145	0.184	0.105
	(0.440)	(0.479)	(0.450)	(0.391)	(0.364)	(0.399)	(0.352)	(0.387)	(0.307)
CSP61	0.225	0.231	0.255	0.117	0.071	0.299	0.186	0.096	0.058
	(0.418)	(0.422)	(0.436)	(0.321)	(0.266)	(0.458)	(0.390)	(0.295)	(0.233)
CSP50	0.151	0.118	0.166	0.279	0.279	0.108	0.203	0.235	0.203
	(0.358)	(0.322)	(0.372)	(0.448)	(0.448)	(0.310)	(0.402)	(0.424)	(0.402)
CSP40	0.112	0.061	0.110	0.173	0.233	0.080	0.258	0.275	0.225
	(0.315)	(0.240)	(0.314)	(0.379)	(0.423)	(0.272)	(0.438)	(0.447)	(0.418)
CSP30	0.043	0.020	0.025	0.053	0.147	0.015	0.057	0.080	0.359
	(0.203)	(0.142)	(0.157)	(0.224)	(0.354)	(0.121)	(0.232)	(0.271)	(0.480)
Number of Observations	71229	26236	12825	3847	3036	16489	3878	2387	2531

DATA APPENDIX TABLE B3

EDP SAMPLE STATISTICS—WOMEN
(Std. Deviations in Parentheses)

Variable Name	Overall	Degree Category 1	2	3	4	5	6	7	8
$DOB_i < 1925$	0.152	0.235	0.206	0.129	0.055	0.034	0.042	0.055	0.056
	(0.359)	(0.424)	(0.405)	(0.336)	(0.229)	(0.181)	(0.202)	(0.228)	(0.230)
$1924 < DOB_i < 1930$	0.047	0.053	0.078	0.045	0.025	0.024	0.017	0.022	0.023
	(0.212)	(0.224)	(0.268)	(0.206)	(0.156)	(0.153)	(0.130)	(0.146)	(0.148)
$1929 < DOB_i < 1935$	0.084	0.096	0.118	0.070	0.043	0.061	0.054	0.049	0.052
	(0.278)	(0.294)	(0.322)	(0.255)	(0.203)	(0.239)	(0.226)	(0.216)	(0.222)
$1934 < DOB_i < 1940$	0.054	0.056	0.069	0.047	0.036	0.050	0.045	0.038	0.047
	(0.226)	(0.229)	(0.254)	(0.211)	(0.185)	(0.218)	(0.208)	(0.190)	(0.212)
$1939 < DOB_i < 1945$	0.093	0.070	0.113	0.086	0.090	0.103	0.108	0.101	0.127
	(0.290)	(0.255)	(0.317)	(0.281)	(0.287)	(0.304)	(0.311)	(0.301)	(0.334)
$1944 < DOB_i < 1950$	0.114	0.077	0.125	0.109	0.116	0.135	0.164	0.156	0.209
	(0.317)	(0.267)	(0.331)	(0.311)	(0.321)	(0.341)	(0.371)	(0.363)	(0.407)
$1949 < DOB_i < 1955$	0.186	0.112	0.180	0.167	0.285	0.247	0.252	0.298	0.354
	(0.389)	(0.315)	(0.384)	(0.373)	(0.451)	(0.431)	(0.434)	(0.457)	(0.478)
$1954 < DOB_i < 1960$	0.120	0.078	0.067	0.178	0.217	0.166	0.169	0.223	0.125
	(0.325)	(0.267)	(0.251)	(0.383)	(0.412)	(0.372)	(0.375)	(0.416)	(0.331)
$1959 < DOB_i < 1977$	0.150	0.224	0.043	0.170	0.133	0.180	0.147	0.059	0.008
	(0.357)	(0.417)	(0.202)	(0.375)	(0.339)	(0.384)	(0.355)	(0.236)	(0.088)
Works in Ile de France	0.254	0.237	0.239	0.286	0.333	0.221	0.316	0.283	0.466
	(0.435)	(0.425)	(0.426)	(0.452)	(0.471)	(0.415)	(0.465)	(0.451)	(0.499)
CSP62	0.227	0.343	0.296	0.108	0.079	0.126	0.073	0.061	0.053
	(0.419)	(0.475)	(0.456)	(0.310)	(0.270)	(0.331)	(0.259)	(0.240)	(0.224)
CSP61	0.050	0.061	0.067	0.027	0.023	0.044	0.027	0.029	0.015
	(0.218)	(0.239)	(0.249)	(0.163)	(0.150)	(0.205)	(0.161)	(0.168)	(0.120)
CSP50	0.458	0.365	0.427	0.596	0.570	0.539	0.630	0.420	0.511
	(0.498)	(0.482)	(0.495)	(0.491)	(0.495)	(0.498)	(0.483)	(0.494)	(0.500)
CSP40	0.073	0.040	0.035	0.090	0.165	0.045	0.097	0.350	0.214
	(0.261)	(0.195)	(0.185)	(0.286)	(0.371)	(0.208)	(0.296)	(0.477)	(0.410)
CSP30	0.013	0.008	0.005	0.016	0.048	0.005	0.009	0.032	0.150
	(0.115)	(0.090)	(0.068)	(0.125)	(0.214)	(0.071)	(0.093)	(0.176)	(0.357)
Number of Observations	57677	19822	12768	4760	3112	10388	2633	3173	1021

DATA APPENDIX TABLE B4

MULTINOMIAL LOGIT ON DEGREE CATEGORIES–MEN
(Std. Errors in Parentheses)

Variable Name	1	2	3	4	5	6	7
Intercept	6.254	5.828	2.465	0.803	3.985	1.714	−0.141
	(0.122)	(0.125)	(0.134)	(0.142)	(0.125)	(0.139)	(0.158)
$1924 < \mathrm{DOB}_i < 1930$	−0.496	−0.320	−0.333	0.005	0.392	0.266	0.102
	(0.105)	(0.106)	(0.131)	(0.133)	(0.113)	(0.132)	(0.179)
$1929 < \mathrm{DOB}_i < 1935$	−0.493	−0.518	−0.344	−0.109	0.734	0.471	0.407
	(0.090)	(0.091)	(0.112)	(0.117)	(0.096)	(0.111)	(0.145)
$1934 < \mathrm{DOB}_i < 1940$	−1.234	−1.117	−0.667	−0.325	0.446	0.318	0.349
	(0.100)	(0.102)	(0.124)	(0.130)	(0.105)	(0.119)	(0.154)
$1939 < \mathrm{DOB}_i < 1945$	−2.031	−1.863	−1.120	−0.381	0.090	0.000	0.519
	(0.085)	(0.087)	(0.105)	(0.106)	(0.089)	(0.102)	(0.126)
$1944 < \mathrm{DOB}_i < 1950$	−2.818	−2.430	−1.307	−0.379	−0.336	−0.216	0.653
	(0.085)	(0.087)	(0.102)	(0.104)	(0.089)	(0.102)	(0.123)
$1949 < \mathrm{DOB}_i < 1955$	−3.388	−3.248	−1.373	−0.069	0.700	−0.363	0.843
	(0.086)	(0.089)	(0.100)	(0.101)	(0.090)	(0.103)	(0.121)
$1954 < \mathrm{DOB}_i < 1960$	−2.289	−2.649	0.074	0.830	0.230	0.312	1.704
	(0.113)	(0.119)	(0.123)	(0.127)	(0.116)	(0.130)	(0.145)
$1959 < \mathrm{DOB}_i < 1977$	1.897	0.246	2.891	2.855	3.319	2.742	3.339
	(0.360)	(0.363)	(0.364)	(0.369)	(0.362)	(0.368)	(0.379)
Unskilled Blue-Collar	−0.850	−1.311	−0.681	−0.193	−1.306	−0.849	−0.155
at Date t in Firm $\mathbf{J}(t,t)$	(0.116)	(0.119)	(0.126)	(0.134)	(0.116)	(0.129)	(0.136)
Skilled Blue-Collar	−0.904	−1.074	−0.557	−0.294	−0.340	−0.006	−0.055
at Date t in Firm $\mathbf{J}(i,t)$	(0.132)	(0.135)	(0.144)	(0.156)	(0.131)	(0.142)	(0.157)
Unskilled White-Collar	−2.758	−2.635	−0.944	−0.217	−2.494	−1.100	−0.437
at Date t in Firm $\mathbf{J}(i,t)$	(0.111)	(0.114)	(0.118)	(0.125)	(0.110)	(0.121)	(0.129)
Skilled White-Collar	−4.028	−3.740	−1.610	−0.377	−3.011	−1.030	−0.100
at Date t in Firm $\mathbf{J}(i,t)$	(0.117)	(0.121)	(0.127)	(0.132)	(0.117)	(0.126)	(0.134)
Manager at Date t	−5.892	−5.996	−3.400	−1.311	−5.195	−3.036	−1.648
in Firm $\mathbf{J}(i,t)$	(0.124)	(0.132)	(0.142)	(0.136)	(0.131)	(0.141)	(0.148)
Works in	−0.627	−0.629	−0.410	−0.265	−0.766	−0.510	−0.399
Ile de France	(0.048)	(0.050)	(0.057)	(0.057)	(0.049)	(0.056)	(0.062)

DATA APPENDIX TABLE B5

MULTINOMIAL LOGIT ON DEGREE CATEGORIES−WOMEN
(Std. Errors in Parentheses)

Variable Name	1	2	3	4	5	6	7
Intercept	7.296	7.148	4.645	2.263	4.555	2.693	2.278
	(0.205)	(0.206)	(0.211)	(0.223)	(0.211)	(0.231)	(0.223)
$1924 < DOB_i < 1930$	−0.723	−0.224	−0.307	0.023	0.391	−0.148	−0.137
	(0.257)	(0.257)	(0.265)	(0.285)	(0.267)	(0.309)	(0.289)
$1929 < DOB_i < 1935$	−0.999	−0.683	−0.742	−0.314	0.441	0.111	−0.201
	(0.199)	(0.200)	(0.207)	(0.225)	(0.208)	(0.233)	(0.224)
$1934 < DOB_i < 1940$	−1.393	−1.073	−1.021	−0.383	0.371	0.054	−0.361
	(0.206)	(0.207)	(0.217)	(0.233)	(0.214)	(0.241)	(0.233)
$1939 < DOB_i < 1945$	−2.328	−1.743	−1.550	−0.542	−0.057	−0.210	−0.439
	(0.169)	(0.169)	(0.177)	(0.189)	(0.177)	(0.199)	(0.189)
$1944 < DOB_i < 1950$	−3.023	−2.429	−2.011	−0.894	−0.529	−0.461	−0.552
	(0.161)	(0.161)	(0.167)	(0.180)	(0.168)	(0.189)	(0.178)
$1949 < DOB_i < 1955$	−3.791	−3.433	−2.537	−0.694	−1.022	−0.927	−0.601
	(0.156)	(0.157)	(0.162)	(0.172)	(0.163)	(0.184)	(0.173)
$1954 < DOB_i < 1960$	−3.082	−3.323	−1.409	0.075	−0.342	−0.264	0.153
	(0.172)	(0.175)	(0.176)	(0.187)	(0.178)	(0.199)	(0.187)
$1959 < DOB_i < 1977$	1.070	−0.673	1.506	2.448	2.753	2.531	1.638
	(0.382)	(0.384)	(0.385)	(0.390)	(0.385)	(0.396)	(0.395)
Unskilled Blue-Collar	−0.205	−0.787	−0.778	−0.248	−0.898	−0.969	−0.511
at Date t in Firm $\mathbf{J}(i,t)$	(0.195)	(0.196)	(0.202)	(0.210)	(0.196)	(0.212)	(0.213)
Skilled Blue-Collar	−0.634	−0.977	−0.840	−0.167	−0.645	−0.675	0.064
at Date t in Firm $\mathbf{J}(i,t)$	(0.295)	(0.296)	(0.308)	(0.320)	(0.297)	(0.320)	(0.315)
Unskilled White-Collar	−2.250	−2.466	−1.218	−0.502	−1.593	−1.008	−0.749
at Date t in Firm $\mathbf{J}(i,t)$	(0.144)	(0.146)	(0.149)	(0.154)	(0.144)	(0.153)	(0.155)
Skilled White-Collar	−3.853	−4.352	−2.379	−0.880	−3.272	−2.062	−0.047
at Date t in Firm $\mathbf{J}(i,t)$	(0.161)	(0.165)	(0.166)	(0.169)	(0.162)	(0.174)	(0.166)
Manager at Date t	−5.449	−6.431	−3.977	−1.725	−5.147	−4.133	−2.052
in Firm $\mathbf{J}(i,t)$	(0.191)	(0.216)	(0.209)	(0.193)	(0.218)	(0.272)	(0.201)
Works in	−0.925	−0.983	−0.738	−0.462	−0.967	−0.541	−0.738
Ile de France	(0.069)	(0.070)	(0.074)	(0.076)	(0.070)	(0.078)	(0.077)

Seniority and Labor Force Experience

In order to impute a level of seniority for left-censored employment spells, we ran regressions (separately for men and women) of seniority on a set of demographic and occupational characteristics using data from the 1978 Salary Structure Survey (ESS, Enquête sur la Structure des Salaires). The results for men are shown in equation (8.1) and the results for women are in (8.2). All regressions included controls for 84 industries.

$$
\begin{aligned}
\text{(8.1)}\quad \text{seniority}_{it} = {} & \underset{(0.081)}{2.513} \\
& + \underset{(0.067)}{14.151}\,[DOB_i \le 1924] + \underset{(0.067)}{12.820}\,[1925 \le DOB_i \le 1929] \\
& + \underset{(0.066)}{10.299}\,[1930 \le DOB_i \le 1934] + \underset{(0.067)}{7.445}\,[1935 \le DOB_i \le 1939] \\
& + \underset{(0.067)}{4.748}\,[1940 \le DOB_i \le 1944] + \underset{(0.065)}{2.569}\,[1945 \le DOBs \le 1949] \\
& + \underset{(0.065)}{0.612}\,[1950 \le DOB_i \le 1954] - \underset{(0.067)}{0.642}\,[1955 \le DOB_i \le 1959] \\
& + \underset{(0.038)}{4.039}\,CSP30_{it} + \underset{(0.031)}{4.939}\,CSP40_{it} \\
& + \underset{(0.037)}{1.885}\,CSP50_{it} + \underset{(0.027)}{2.898}\,CSP61_{it} \\
& - \underset{(0.026)}{0.958}\,\text{Ile de France}_{it},
\end{aligned}
$$

$N = 547{,}746, \qquad R^2 = 0.461,$

$$
\begin{aligned}
\text{(8.2)}\quad \text{seniority}_{it} = {} & \underset{(0.084)}{2.114} \\
& + \underset{(0.074)}{12.669}\,[DOB_i \le 1924] + \underset{(0.075)}{11.014}\,[1925 \le DOB_i \le 1929] \\
& + \underset{(0.073)}{8.979}\,[1930 \le DOB_i \le 1934] + \underset{(0.074)}{7.278}\,[1935 \le DOB_i \le 1939] \\
& + \underset{(0.075)}{5.989}\,[1940 \le DOB_i \le 1944] + \underset{(0.070)}{4.604}\,[1945 \le DOB_i \le 1949] \\
& + \underset{(0.068)}{2.822}\,[1950 \le DOB_i \le 1954] + \underset{(0.068)}{0.641}\,[1955 \le DOB_i \le 1959] \\
& + \underset{(0.082)}{5.116}\,CSP30_{it} + \underset{(0.057)}{5.789}\,CSP40_{it} \\
& + \underset{(0.037)}{1.442}\,CSP50_{it} + \underset{(0.054)}{2.429}\,CSP61_{it} \\
& - \underset{(0.031)}{0.988}\,\text{Ile de France}_{it},
\end{aligned}
$$

$N = 260{,}580, \qquad R^2 = 0.373,$

where

$$
\begin{aligned}
&DOB_i = \text{Date of birth of individual } i,\\
&CSP30_{it} = 1 \text{ if } i \text{ is an engineer, professional, or manager},\\
&CSP40_{it} = 1 \text{ if } i \text{ is technician or technical white-collar},\\
&CSP50_{it} = 1 \text{ if } i \text{ is any other white-collar},\\
&CSP61_{it} = 1 \text{ if } i \text{ is a skilled blue-collar},\\
&CSP62_{it} = 1 \text{ if } i \text{ is an unskilled blue-collar (omitted)},\\
&\text{Ile de France}_{it} = 1 \text{ if the establishment is in Ile-de-France}.
\end{aligned}
\tag{8.3}
$$

The excluded date of birth category was $1960 \leq DOB_i$. The coefficients on the industry indicators are not shown.

To compute the values of seniority and labor force experience, we used the following algorithms. If the individual was left-censored and the imputed job seniority was negative, we set job seniority prior to 1976 to zero. If the individual was first observed after 1976, we assumed that job seniority on that job prior to the date of the first DAS observation for the individual was zero. If the age at the date of any observation (1976 or otherwise) was less than the expected school-leaving age, both total labor force experience and prior job seniority were set to zero. In all other cases (when the age was greater than the expected school-leaving age), we calculated total labor market experience and job seniority as follows. If the observation was the earliest appearance of the individual in our data, we set job seniority equal to job seniority up to the date of the first observation plus the number of days worked for that enterprise in the year of the first observation, divided by 360 and we set total labor market experience to the current age less the school-leaving age. If the observation was not the first for the individual but there was an observation in the previous year for the person,[44] we added 1 to total labor market experience. If the individual was employed for the majority of the current year by the same enterprise that employed him or her for the majority of the previous year, i.e. $SIREN_t = SIREN_{t-1}$, we added 1 to the level of seniority at $t-1$. If $SIREN_t \neq SIREN_{t-1}$, we set seniority equal to the number of days worked divided by 360.

If, on the other hand, there was no observation in the previous year, we distinguished between $t = 1982$ or $t = 1984$ and other years. When $t \neq 1982$ or 1984, total labor market experience was increased by 1 (reflecting experience gained in the year of the observation). If the current *SIREN* and the most recent previous *SIREN* were the same, we added the number of days worked divided by 360 to the most recent previous level of seniority. This is similar to assuming that the worker was temporarily laid off but retained his or her seniority in the firm when recalled. Otherwise, we set seniority to the number of days worked divided by 360.

In the case where $t = 1982$ or $t = 1984$, if the preceding observation was 2 years earlier (i.e. the missing data only occurred over a period when no data were available for any individual), we increased total labor market experience by 2. If $SIREN_{t-2} = SIREN_t$, seniority was increased by 2. If $SIREN_{t-2} \neq SIREN_t$, seniority was increased by 0.5 plus the number of days worked divided by 360.[45]

[44] The structure of our database is such that this condition (observations for individual i at both t and $t-1$) could only fail to be satisfied under 3 conditions. The first is that the individual was employed in the civil service in the intervening years. The second is that the individual was unemployed for an entire calendar year. The third is that $t = 1982$ or $t = 1984$, since we were not given access to the data for 1981 or 1983. We largely discount the first possibility for the reasons mentioned in the text. The other two possibilities are treated explicitly.

[45] We assumed that the probability the individual was reemployed in the missing year was equal to the probability that the individual was reemployed in the observation year. Thus, the expected increment to job seniority is the share of the year worked in the observation year plus $(\frac{1}{2} \cdot 0) + (\frac{1}{2} \cdot 1) = 0.5$.

If the preceding observation was more than 2 years earlier, we increased total labor market experience by 1.5.[46] If the current *SIREN* and the most recent previous *SIREN* were the same, we added the number of days worked divided by 360 plus 0.5 to the most recent previous level of seniority. This is similar to assuming that the worker was recalled from temporary layoff with equal probability in the observation year and in the missing year. If the two *SIREN*s were different, we set seniority to 0.5 plus the number of days worked divided by 360.

Elimination of Outliers

We ran a standard log earnings regression (the dependent variable was the logarithm of real annualized compensation cost, *LFRAISRE*, the same one used in the analyses reported in Tables II–XII) on our DAS data and considered all observations that were more than 5 standard deviations away from their predicted values as outliers. These observations were discarded. The estimated coefficients of this earnings regression are shown in equation (8.4).

$$
\begin{aligned}
(8.4)\qquad LFRAISRE_{it} = &-\underset{(0.005)}{3.250} \\
&+\underset{(0.000)}{0.210}\ \text{Male}_i &&+\underset{(0.000)}{0.123}\ \text{Ile de France}_{it} \\
&+\underset{(0.000)}{0.082}\ \text{Year}_{\text{it}} &&+\underset{(0.002)}{0.056}\ \text{Degree Category } 2_i \\
&+\underset{(0.002)}{0.415}\ \text{Degree Category } 3_i &&+\underset{(0.003)}{0.627}\ \text{Degree Category } 4_i \\
&+\underset{(0.001)}{0.266}\ \text{Degree Category } 5_i &&+\underset{(0.003)}{0.642}\ \text{Degree Category } 6_i \\
&+\underset{(0.002)}{0.648}\ \text{Degree Category } 7_i &&+\underset{(0.003)}{1.421}\ \text{Degree Category } 8_i \\
&+\underset{(0.000)}{0.055}\ \text{Experience}_{it} &&-\underset{(0.003)}{0.222}\ \text{Experience}^2_{it} \\
&+\underset{(0.001)}{0.052}\ \text{Experience}^3_{it} &&-\underset{(0.000)}{0.005}\ \text{Experience}^4_{it},
\end{aligned}
$$

$N = 5{,}325{,}352, \quad R^2 = 0.437, \quad \sigma = 0.477.$

Definition of Z Variables and Coefficients in the Conditional Method

Data Appendix Table B6 contains the definitions, regression coefficients, and coefficient standard errors for the Z variables used in estimating the statistical model (3.17) as reported in Table III in the column labelled "Conditional Method Persons First."

Pooled Regression for Order-Dependent Persons-First Estimation

Recovery of the firm effects was done in the conditional methods on a firm-by-firm basis. All observations corresponding to firms for which there were fewer than 10 observations were grouped together and included in a single, pooled regression. The results of this regression for the pooled "firm" in the order-dependent, persons-first case are shown in equation (8.5). The results for the order-independent pooled "firm" are not shown.

$$
(8.5)\qquad DLFRAISR_{it} = -\underset{(3.375e\text{-}4)}{0.028} + \underset{(8.476e\text{-}5)}{0.003}\ s_{it} - \underset{(1.772e\text{-}4)}{0.005}\ \mathbf{T}_1(s_{it}-10),
$$

$N = 1{,}353{,}794, \quad R^2 = 0.0013.$

[46] We assumed that the probability the individual was reemployed in the missing year was equal to the probability that the individual was reemployed in the observation year. Thus, the expected increment to total labor market experience is $(\frac{1}{2}\cdot 1) + (\frac{1}{2}\cdot 2) = 1.5$.

DATA APPENDIX TABLE B6

SUMMARY STATISTICS, COEFFICIENTS AND STANDARD ERRORS FOR Z VARIABLES IN THE CONDITIONAL METHOD ORDER INDEPENDENT ESTIMATION

Variable Definition	Mean	Standard Deviation	Coefficient	Standard Error
Firm size × average experience	2.54E-05	3.16E-06	1.11E-05	3.82E-06
Firm size × age at end of school	1.79E-04	2.91E-06	1.77E-05	3.58E-06
Firm size squared × average experience	− 7.57E-08	2.00E-08	6.38E-08	2.00E-08
Firm size squared × age at end of school	− 5.28E-07	1.00E-08	− 3.06E-08	2.00E-08
Firm size × seniority × average experience	3.23E-06	3.30E-07	2.76E-06	3.50E-07
Firm size × seniority × age at end of school	− 1.43E-05	4.50E-07	− 6.95E-06	4.30E-07
Firm size squared × seniority × average experience	− 5.76E-09	1.60E-07	− 1.12E-09	7.19E-10
Firm size squared × seniority × age at end of school	4.47E-08	1.00E-08	2.01E-08	1.91E-07
Industry 1 × average experience	− 3.92E-04	1.28E-04	− 2.06E-03	3.19E-03
Industry 1 × age at end of school	− 2.22E-02	1.48E-04	1.04E-02	3.49E-03
Industry 1 × seniority × average experience	3.95E-04	1.61E-05	1.50E-04	1.53E-05
Industry 1 × seniority × age at end of school	− 2.98E-04	2.51E-05	− 1.12E-04	2.17E-05
Industry 2 × average experience	2.10E-03	2.17E-04	− 4.71E-03	3.20E-03
Industry 2 × age at end of school	1.62E-02	2.20E-04	1.39E-02	3.50E-03
Industry 2 × seniority × average experience	− 1.25E-04	2.18E-05	− 1.41E-04	2.27E-05
Industry 2 × seniority × age at end of school	6.14E-04	3.14E-05	3.32E-04	3.04E-05
Industry 3 × average experience	3.82E-04	7.75E-05	− 1.93E-03	3.19E-03
Industry 3 × age at end of school	− 3.61E-02	8.33E-05	1.03E-02	3.49E-03
Industry 3 × seniority × average experience	2.07E-04	8.98E-06	8.41E-05	8.00E-06
Industry 3 × seniority × age at end of school	− 4.80E-05	1.36E-05	− 1.52E-05	1.14E-05
Industry 4 × average experience	− 2.52E-04	7.46E-05	− 2.15E-03	3.19E-03
Industry 4 × age at end of school	− 1.76E-02	7.08E-05	1.08E-02	3.49E-03
Industry 4 × seniority × average experience	4.09E-05	8.03E-06	8.92E-05	7.62E-06
Industry 4 × seniority × age at end of school	3.66E-04	1.12E-05	− 1.38E-05	9.93E-06
Industry 5 × average experience	2.16E-03	8.12E-05	− 1.95E-03	3.19E-03
Industry 5 × age at end of school	− 3.59E-02	8.61E-05	9.18E-03	3.49E-03
Industry 5 × seniority × average experience	2.92E-04	9.78E-06	1.14E-04	9.37E-06
Industry 5 × seniority × age at end of school	− 4.70E-04	1.48E-05	7.38E-06	1.29E-05
Industry 6 × average experience	1.02E-03	8.46E-05	1.67E-04	3.19E-03
Industry 6 × age at end of school	− 2.94E-02	1.05E-04	4.62E-03	3.49E-03
Industry 6 × seniority × average experience	7.20E-04	1.20E-05	1.07E-04	1.07E-05
Industry 6 × seniority × age at end of school	− 1.41E-03	1.85E-05	− 1.00E-04	1.50E-05
Industry 7 × average experience	− 3.46E-04	6.93E-05	− 2.16E-03	3.19E-03
Industry 7 × age at end of school	6.53E-03	8.00E-05	8.91E-03	3.49E-03
Industry 7 × seniority × average experience	− 4.73E-05	9.20E-06	− 4.40E-05	8.55E-06
Industry 7 × seniority × age at end of school	9.89E-04	1.38E-05	2.34E-04	1.17E-05
Industry 8 × average experience	− 3.60E-04	1.32E-04	− 2.88E-03	3.19E-03
Industry 8 × age at end of school	2.35E-02	1.39E-04	9.77E-03	3.49E-03
Industry 8 × seniority × average experience	− 3.22E-04	1.53E-05	7.68E-05	1.49E-05
Industry 8 × seniority × age at end of school	1.70E-03	2.25E-05	1.02E-04	2.06E-05
Industry 9 × average experience	5.22E-04	5.53E-05	− 2.81E-03	3.19E-03
Industry 9 × age at end of school	3.57E-02	5.89E-05	8.25E-03	3.49E-03
Industry 9 × seniority × average experience	− 3.87E-04	6.81E-06	− 2.85E-05	6.40E-06
Industry 9 × seniority × age at end of school	1.89E-03	9.63E-06	1.79E-04	8.36E-06
Industry 10 × average experience	− 1.98E-03	8.29E-05	− 3.20E-03	3.19E-03
Industry 10 × age at end of school	3.43E-02	7.92E-05	8.87E-03	3.49E-03
Industry 10 × seniority × average experience	− 1.10E-04	9.56E-06	− 1.97E-05	1.01E-05
Industry 10 × seniority × age at end of school	0.001673	0.00001264	0.000238	0.00001243

Notes: These coefficients supplement the coefficients reported in Table III, Column "Conditional Method Persons First."

DATA APPENDIX TABLE B7

DESCRIPTIVE STATISTICS FOR BASIC INDIVIDUAL LEVEL VARIABLES BY SEX FOR 1976 TO 1987

	Men		*Women*	
Variable Definition	Mean	Standard Deviation	Mean	Standard Deviation
Real Total Annual Compensation Cost, 1,000FF 1980	89.0967	61.6302	67.3646	37.4208
Log (Real Annual Compensation Cost, 1980 FF)	4.3442	0.5187	4.0984	0.4801
Total Labor Force Experience	17.2531	11.8258	15.4301	12.0089
(Total Labor Force Experience)^2/100	4.3752	4.9197	3.8230	4.9440
(Total Labor Force Experience)^3/1,000	13.1530	19.4305	11.6079	19.6863
(Total Labor Force Experience)^4/10,000	43.3453	77.9542	39.0589	80.3251
Seniority	7.7067	7.5510	6.5437	6.5268
Lives in Ile-de-France (Paris Metropolitan Region)	0.2561		0.2910	
No Known Degree	0.3064	0.2190	0.2971	0.2124
Completed Elementary School	0.1556	0.1458	0.1893	0.1739
Completed Junior High School	0.0565	0.0792	0.0869	0.1008
Completed High School (Baccalauréat)	0.0528	0.0804	0.0711	0.0881
Basic Vocational-Technical Degree	0.2652	0.1849	0.1926	0.1545
Advanced Vocational-Technical Degree	0.0701	0.0893	0.0532	0.0802
Technical College or University Diploma	0.0469	0.0754	0.0838	0.1247
Graduate School Diploma	0.0465	0.0964	0.0259	0.0551
Year of data	81.3106	3.7250	81.4730	3.7180
Number of Observations for the Firm in Sample	4,402.3800	16,164.6200	1,605.3100	7,797.1300
Observations	3,434,530		1,870,578	
Persons	711,518		454,787	
Proportion with Identified Least Squares Estimate of Individual and Firm Effect	0.7425		0.7448	

Source: Authors' calculations based on the Déclarations annuelles des salaires (DAS).

Construction of the Operating Income Variable

The operating income variable (excedent brut d'exploitation) was constructed as in the following equation:

(8.6) *EBE* = ventes de marchandises (merchandise sold)
− achat de marchandises (merchandise purchased)
− variation de stock de marchandises (variation in merchandise inventory)
+ ventes de biens (goods sold)
+ ventes de services (services sold)
+ production stockée (inventoried production)
+ production immobilisée (unfinished production)
− achats de matières premières (primary materials purchased)
− variation de stocks sur matières premières (variation of primary materials inventories)

− autres achats et charges externes

(other purchases and outside charges)

+ subventions d'exploitation (incentives for production)

− impôts, taxes et versements assimilés

(value added tax and other accrued taxes on

or credits for production)

− salaires et traitements (salaries and benefits)

− charges sociales (payroll taxes).

REFERENCES

ABOWD, JOHN M. (1989): "The Effect of Wage Bargains on the Stock Market Value of the Firm," *American Economic Review*, 79, 774–800.

——— (1990): "Does Performance-Based Managerial Compensation Affect Corporate Performance," *Industrial and Labor Relations Review*, 43, S52–S73.

ABOWD, JOHN M., AND LAURENCE ALLAIN (1996): "Compensation Structure and Product Market Competition," *Annales d'Économie et de Statistique*, 41/42, 207–217.

ABOWD, JOHN M., AND M. BOGNANNO (1995): "International Differences in Executive and Managerial Compensation," in *Differences and Changes in Wage Structures*, ed. by R. B. Freeman and L. Katz. Chicago: University of Chicago Press for the NBER, 67–103.

ABOWD, JOHN M., AND DAVID CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.

ABOWD, JOHN M., JOHN HANNON, AND GEORGE MILKOVICH (1990): "The Effects of Human Resource Management Decisions on Shareholder Value," *Industrial and Labor Relations Review*, 43, S203–S236.

ABOWD, JOHN M., AND FRANCIS KRAMARZ (1993): "A Test of Negotiation and Incentive Compensation Models Using Longitudinal French Enterprise Data," in *Labour Demand and Equilibrium Wage Formation*, ed. by J. C. van Ours, G. A. Pfann, and G. Ridder. Amsterdam: North-Holland, pp. 111–146.

ABOWD, JOHN M., FRANCIS KRAMARZ, THOMAS LEMIEUX, AND DAVID N. MARGOLIS (forthcoming): "Minimum Wages and Youth Employment in France and the United States," in *Youth Unemployment and Employment in Advanced Countries*, ed. by D. Blanchflower and R. Freeman. Chicago: University of Chicago Press for the NBER.

ABOWD, JOHN M., FRANCIS KRAMARZ, DAVID N. MARGOLIS, AND KENNETH TROSKE (1998): "The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States," Cornell University Working Paper.

ABOWD, JOHN M., FRANCIS KRAMARZ, AND ANTOINE MOREAU (1996): "Product Quality and Worker Quality," *Annales d'Économie et de Statistique*, 41/42, 299–322.

ABOWD, JOHN M., AND THOMAS LEMIEUX (1993): "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada," *Quarterly Journal of Economics*, 108, 983–1014.

ABRAHAM, KATHARINE G., AND HENRY S. FARBER (1987): "Job Duration, Seniority and Earnings," *American Economic Review*, 77, 278–297.

ALTONJI, JOSEPH, AND R. A. SHAKOTKO (1987): "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54, 437–459.

BECKER, GARY S. (1993): *Human Capital*, 3rd Edition. Chicago: University of Chicago Press.

BECKER, GARY S., AND GEORGE J. STIGLER (1974): "Law Enforcement, Malfeasance and Compensation of Enforcers," *Journal of Legal Studies*, 3:1, 1–18.

BELZIL, CHRISTIAN (1996): "Employment Reallocation, the Return to Human Capital and the Allocation of Workers between Expanding and Declining Firms," Concordia University, Department of Economics, Working Paper.

BINGLEY, PAUL, AND NIELS WESTERGÅRD-NIELSEN (1996): "Individual Wages within and between Establishments," University of Århus, Working Paper.

BLANCHFLOWER, DAVID G., ANDREW J. OSWALD, AND PETER SANFEY (1996): "Wages, Profits, and Rent-Sharing," *Quarterly Journal of Economics*, 111, 227–251.

BLAU, FRANCINE, AND LAWRENCE KAHN (1996): "International Differences in Male Wage Inequality: Institutions vs. Market Forces," *Journal of Political Economy*, 104, 791–836.

BROWN, CHARLES, AND JAMES L. MEDOFF (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.

BROWN, JAMES N. (1989): "Why Do Wages Rise with Seniority?" *American Economic Review*, 79, 971–991.

BROWN, JAMES N., AND ORLEY ASHENFELTER (1986): "Testing the Efficiency of Employment Contracts," *Journal of Political Economy*, 94, S40–S87.

BULOW, JEREMY, AND LAWRENCE SUMMERS (1976): "A Theory of Dual Labor Markets with Applications to Industrial Policy, Discrimination and Keynesian Unemployment," *Journal of Labor Economics*, 4, 376–414.

CAHUC, PIERRE, AND BRIGITTE DORMONT (1997): "Profit Sharing: Does It Increase Productivity and Employment? A Theoretical Model and Empirical Evidence on French Micro Data," *Labour Economics*, 4, 293–319.

CAHUC, PIERRE, AND FRANCIS KRAMARZ (1997): "Voice and Loyalty as a Delegation of Authority: A Model and a Test on a Matched Worker-Firm Panel," *Journal of Labor Economics*, 15, 658–688.

CAIN, GLENN (1976): "The Challenge of Segmented Labor Market Theories to Orthodox Theory," *Journal of Economic Literature*, 14, 1215–1257.

CARD, DAVID (1986): "Efficient Contracts and Costs of Adjustment: Short Run Employment Determination for Airline Mechanics," *American Economic Review*, 76, 1045–1071.

CARD, DAVID, FRANCIS KRAMARZ, AND THOMAS LEMIEUX (1996): "Changes in the Relative Structure of Wages and Employment: A Comparison of the United States, Canada, and France," NBER Working Paper No. 5487.

CEREQ-DEP-INSEE (1990): *Bilan Formation-Emploi*. Paris: République Française.

CHAMBERLAIN, GARY (1984): "Panel Data," in *Handbook of Econometrics*, ed. by Z. Griliches and M. D. Intrilligator. Amsterdam: North Holland.

CHRISTOFIDES, LOUIS, AND ANDREW J. OSWALD (1991): "Efficient and Inefficient Employment Contracts A Study Based on Canadian Contracts," *Research in Labor Economics*, 12, 173–190.

——— (1992): "Real Wage Determination and Rent-Sharing in Collective Bargaining Agreements," *Quarterly Journal of Economics*, 107, 985–1002.

DAVIS, STEVEN J., AND JOHN HALTIWANGER (1996): "Employer Size and the Wage Structure in the United States," *Annale d'Économie et de Statistique*, 41/42, 323–368.

DICKENS, WILLIAM T., AND LAWRENCE F. KATZ (1987): "Inter-Industry Wage Differences and Industry Characteristics," in *Unemployment and the Structure of Labor Markets*, ed. by Kevin Lang and Jonathan S. Leonard. Oxford: Basil Blackwell.

DICKENS, WILLIAM T., AND KEVIN LANG (1985): "A Test of Dual Labor Market Theory," *American Economic Review*, 75, 792–805.

EHRENBERG, RONALD G., EDITOR (1990): *Do Compensation Policies Matter?* Ithaca, NY: ILR Press.

EHRENBERG, RONALD G., AND GEORGE T. MILKOVICH (1987): "Compensation and Firm Performance," in *Human Resources and the Performance of the Firm*, ed. by S. G. Allen et al. Madison, WI: IRRA, pp. 87–122.

ENTORF, HORST, MICHEL GOLLAC, AND FRANCIS KRAMARZ (forthcoming): "New Technologies, Wages and Worker Selection," forthcoming in the *Journal of Labor Economics*.

ENTORF, HORST, AND FRANCIS KRAMARZ (1997): "Does Unmeasured Ability Explain the Higher Wage of New Technology Workers," *European Economic Review*, 41, 1489–1510.

——— (forthcoming): "The Impact of New Technologies on Wages: Lessons from Matching Panels on Employees and on Their Firms," *Economics of Innovation and News Technology*.

FLINN, CHRISTOPHER (1986): "Wages and Job Mobility of Young Workers," *Journal of Political Economy*, 94, S88–S110.

GIBBONS, ROBERT, AND LAWRENCE KATZ (1991): "Layoffs and Lemons," *Journal of Labor Economics*, 9, 351–380.

——— (1992): "Does Unmeasured Ability Explain Inter-Industry Wage Differentials'" *Review of Economic Studies*, 59, 515–535.

GIBBONS, ROBERT, AND KEVIN J. MURPHY (1990): "Relative Performance Evaluation for Chief Executive Officers," *Industrial and Labor Relations Review*, 43, S30–S51.

——— (1992): "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100, 468–505.

GOUX, DOMINIQUE, AND ERIC MAURIN (forthcoming): "Persistence of Inter-Industry Wage Differentials: A Re-examination on Matched Worker-firm Panel Data," forthcoming in the *Journal of Labor Economics*.

GROSHEN, ERICA (1991): "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?" *Quarterly Journal of Economics*, 106, 869–884.

——— (1996): "American Employer Salary Surveys and Labor Economics Research: Issues and Contributions," *Annales d'Économie et de Statistique*, 41/42, 413–442.

HART, OLIVER, AND BENGT HOLMSTRÖM (1987): "A Theory of Contracts," in *Advances in Economic Theory*, *Fifth World Congress*, ed. by Truman F. Bewley. Cambridge: Cambridge University Press, pp. 71–155.

HAUSMAN, JERRY A. (1978): "Specification Tests in Econometrics," *Econometrica*, 546, 1251–1271.

HECKMAN, JAMES J., AND G. SEDLACEK (1985): "Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93, 1077–1125.

HUTCHENS, ROBERT (1987): "A Test of Lazear's Theory of Delayed Payment Contracts," *Journal of Labor Economics*, 5, S153–S170.

ICHNIOWSKI, CASEY, AND KATHRYN SHAW (1993): "Do High Performance Work Systems Really Work?" Columbia University Working Paper.

INSEE (1989): *Echantillon d'Entreprises État d'Avancement*. Paris: République Française, No. 13/G231.

——— (1990a): *Echantillon d'Entreprises Fichier de Réference de l'Échantillon*. Paris: Republique Française, No. 06/G231.

——— (1990b): *Echantillon d'Entreprises Mise à Disposition des Données BIC*. Paris: République Française, No. 10/G231.

——— (1990c): *Echantillon d'Entreprises Mise à Disposition des Données ESE*. Paris: République Française, No. 64/G231.

JENSEN, MICHAEL C., AND KEVIN J. MURPHY (1990): "Performance Pay and Top Management Incentives," *Journal of Political Economy*, 98, 225–264.

JOVANOVIC, BOYAN (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972–990.

JUHN, CHINHUI, KEVEN M. MURPHY, AND BROOKS PIERCE (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101, 410–442.

KAHN, LAWRENCE, AND PETER SHERER (1990): "Contingent Pay and Managerial Performance," *Industrial and Labor Relations Review*, 43, S52–S73.

KRAMARZ, FRANCIS, STÉFAN LOLLIVIER, AND LOUIS-PAUL PELÉ (1996): "Wage Inequalities and Firm-Specific Compensation Policies in France," *Annales d'Economie et de Statistique*, 41/42, 369–386.

KRAMARZ, FRANCIS, AND PATRICK REY (1995): "The Dynamics of Wages," INSEE Département de la Recherche, Working Paper.

KRUEGER, ALAN B. (1993): "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984–1989," *Quarterly Journal of Economics*, 108, 33–60.

Krueger, Alan B., and Lawrence H. Summers (1987): "Reflections on the Interindustry Wage Structure," in *Unemployment and the Structure of Labor Markets*, ed. by Kevin Lang and Jonathan S. Leonard. New York: Basil Blackwell.

——— (1988): "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 259–293.

Lazear, Edward P. (1979): "Why Is There Mandatory Retirement?" *Journal of Political Economy*, 87, 1261–1284.

——— (1998): *Personnel Economics for Managers*. New York: John Wiley and Sons.

Leonard, Jonathan S. (1990): "Executive Pay and Firm Performance," *Industrial and Labor Relations Review*, 43, S13–S29.

Leonard, Jonathan S., and Marc Van Audenrode (1996a): "Worker's Limited Liability, Turnover and Employment Contracts," *Annales d'Économie et de Statistique*, 41/42, 41–78.

——— (1996b): "Persistence of Firm and Wage Components," University of California at Berkeley Working Paper.

Leontief, Wassily W. (1946): "The Pure Theory of the Guaranteed Annual Wage Contract," *Journal of Political Economy*, 54, 76–79.

MacDonald, Ian, and Robert M. Solow (1981): "Wage Bargaining and Employment," *American Economic Review*, 71, 886–908.

MaCurdy, Thomas, and John Pencavel (1986): "Testing Between Competing Models of Wage and Employment Determination in Unionized Markets," *Journal of Political Economy*, 94, S3–S39.

Manning, Alan (1987): "An Integration of Trade Union Models in a Sequential Bargaining Framework," *Economic Journal*, 97, 121–139.

Margolis, David N. (1993): "Compensation Practices and Government Policies in Western European Labor Markets," Ph.D. Dissertation, Cornell University.

——— (1996): "Cohort Effects and Returns to Seniority in France," *Annales d'Économie et de Statistique*, 41/42, 443–464.

Meron, Monique (1988): "Les Migrations des Salariés de l'État: Plus Loin de Paris, plus Près du Soleil," *Economie et Statistique*, 214, 3–18.

Mundlak, Yair (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85.

Murphy, Kevin M., and Robert H. Topel (1987): "Unemployment, Risk, and Earnings: Testing for Equalizing Wage Differences in the Labor Market," in *Unemployment and the Structure of the Labor Market*, ed. by Kevin Lang and Jonathan S. Leonard. New York: Basil Blackwell.

Murphy, Kevin M., and Finis Welch (1992): "The Structure of Wages," *Quarterly Journal of Economics*, 107, 285–326.

Nickell, Stephen, and Sushil Wadhwani (1991): "Employment Determination in British Industry: Investigations Using Micro-Data," *Review of Economic Studies*, 58, 955–970.

Pagan, Adrian (1984): "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25, 21–47.

Rosen, Sherwin (1986): "The Theory of Equalizing Differences," in *Handbook of Labor Economics*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam: North Holland, pp. 641–692.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economics Papers*, 3, 35–146.

Rubin, Donald (1987): *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Sappington, David (1991): "Incentives in Principal-Agent Relationships," *Journal of Economic Perspectives* 5, 45–66.

Scheffé, H. (1959): *The Analysis of Variance*. New York: John Wiley and Sons.

Searle, Shayle R., George Casella, and Charles E. McCulloch (1992): *Variance Components*. New York: John Wiley and Sons.

Shapiro, Carl, and Joseph E. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433–444.

Thaler, Richard H. (1989): "Anomalies: Interindustry Wage Differentials," *Journal of Economic Perspectives*, 3, 181–193.

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

Topel, Robert, and Michael Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107, 441–479.

Weiss, Andrew, and Henry J. Landau (1984): "Wages, Hiring Standards, and Firm Size," *Journal of Labor Economics*, 2, 477–499.

Willis, Robert (1986): "Wage Determinants: A Survey," in *Handbook of Labor Economics*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam: North Holland, pp. 525–602.