ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 41/42 — 1996

Cohort Effects and Returns to Seniority in France

David N. MARGOLIS*

ABSTRACT. – We examine empirically returns to seniority in France and estimate cohort effects in both firm specific compensation policies and returns to job seniority. We demonstrate the biases in several estimators of returns to seniority and show that allowing firm specific compensation policies and returns to seniority to vary by entry cohort uncovers patterns of variance within firm and within cohort in estimated intercepts and slopes that can be explained by no single theory.

Les effets de cohorte et le rendement d'ancienneté en France

RÉSUMÉ. – Nous examinons les rendements d'ancienneté en France, et nous estimons les effets de cohorte à la fois dans la politique de rémunération de l'entreprise et dans son rendement d'ancienneté. Nous soulignons les biais présents dans plusieurs estimateurs de rendement d'ancienneté, et nous montrons que la flexibilité introduite par l'estimation des effets de cohorte nous permet de decouvrir des variations à travers les cohortes et à travers les entreprises qui ne sont cohérentes avec aucun des théories de rémunération trouvées dans la littérature.

* D. N. MARGOLIS: Université de Montréal-CIRANO-CRDE, Département de Sciences Économiques. The author would like to thank John ABOWD and Kenneth BURDETT for helpful discussions. The author would especially like to thank Francis KRAMARZ for his input. An earlier version of this paper was prepared for the conference "The Microeconometrics of Human-Resource Management: Multinational Studies of Firm Practices", Paris, France, December 7-9, 1994. This research was supported by the National Science Foundation (SBR 91-11186 and SBR 93-21053) and the CAFIR program at the Université de Montréal.

1 Introduction

Many people have estimated returns to job seniority for the United States [1], and the debate remains open as to whether such returns exist empirically. Even on the theoretical plane, the regularity and size of within-job increases in wages over-and-above market increases is still far from being an issue of universal consensus [2]. This paper proposes to inform this debate by empirically examining returns to seniority in France [3] and by explicitly estimating cohort effects in returns to seniority [4]. The question of cohort effects in returns to seniority is potentially important, as the behavior of these effects can serve as a way to distinguish between competing theoretical models of returns to seniority [5]. Here, we simply demonstrate the value of being able to distinguish between entering cohorts when estimating firm-specific returns to seniority.

We find that, according to the econometric technique used, the estimates of returns to job seniority in France vary dramatically. The most basic estimator (ordinary least squares) suggests returns to seniority of 6-7 percent per year. Accounting for unobserved individual heterogeneity by fixed effects causes this estimate to drop to around 3 percent per year. Allowing for heterogeneity in firm compensation policies via the TOPEL [1991] technique causes estimated returns to seniority to rise back up to around 4.5 percent per year. Allowing still more heterogeneity, this time in both compensation policy and seniority returns [6] (via the projection method of ABOWD, KRAMARZ and MARGOLIS [1994]) implies zero mean returns to seniority across firms as well as across individuals, although the standard deviation of the point estimates of returns to seniority is large enough to encompass even the

1. *See*, for example, ABRAHAM and FARBER [1987], ABRAHAM and FARBER [1988], ALTONJI and SHAKOTKO [1987] and TOPEL [1991].

2. *See*, for example, BECKER [1993], BEAUDRY and DINARDO [1991], HARRIS and HOLMSTROM [1982], JOVANOVIC [1979], and MACLEOD and MALCOMSON [1988].

3. Other attempts to estimate returns to seniority in France include ABOWD, KRAMARZ and MARGOLIS [1994] and MARGOLIS [1995b].

4. BAKER, GIBBS and HOLMSTROM [1994] also estimate cohort effects on wages and show that there is significant variation within a firm across cohorts in wages for workers with equivalent other characteristics. Unfortunately, their analysis makes use of data from only one firm, and thus it is not clear how general the phenomenon of cohort effects might be.

5. For a more detailed examination of the implications of various theoretical models for cohort specific returns to job seniority, *see* MARGOLIS [1995a].

6. We use the term "compensation policy" to designate the starting compensation paid by a given firm, and we use returns to seniority to describe the evolution of compensation within a firm after hiring.

OLS results within a one-standard deviation band around the mean. Finally, when we allow firm specific compensation policies and firm specific returns to seniority to vary by entry cohort, we again find zero mean returns, but we find significant variance in both the compensation policy and returns to seniority. This variance is largest across firm/within cohort for compensation policy. Such variation suggests that one needs to take care to account for cohort and firm heterogeneity when modelling returns to seniority, both in theoretical and empirical work, in particular because there exists no single theory that can explain the patterns of variance within firm and within cohort that we observe in our data [7].

The structure of the paper is as follows. In section 2 we lay out the statistical model that we use as a framework for interpreting the different estimators and in section 3 we briefly describe the data we will use in the estimations. We then estimate returns to job seniority by Ordinary Least Squares (section 4) and by OLS allowing for worker-specific fixed effects (section 5). We then apply the estimation strategy suggested by TOPEL [1991] in section 6 and the strategy suggested by ABOWD, KRAMARZ and MARGOLIS [1994] in section 7. We finally apply a new estimation method, based on the approach suggested by ABOWD, KRAMARZ and MARGOLIS [1994], to estimate different seniority returns for different cohorts within a given firm in section 8. At each step along the estimation path, which basically traces the evolution of thought with regard to important concerns in estimating returns to seniority, we discuss the sorts of biases in the estimated returns that are eliminated by the particular method, as well as those that remain. We then suggest an interpretation for the results of the cohort-specific estimations in the context of several popular explanations of seniority returns in section 9. Section 10 concludes.

2 An Encompassing Descriptive Model of Returns to Job Seniority

In order to be able to incorporate the implications of the variety of different explanations for returns to job seniority into a single theoretical model, we need to specify a very general empirical model of wage determination. In this paper, we estimate variants on the following specification.

$$(1)\quad w_{i,T_0+t} = X_{i,T_0+t}\beta + \alpha_i + \phi_{J(i,T_0+t),T_0} + \gamma_{J(i,T_0+t),T_0} s_{i,T_0+t} + \eta_{i,T_0+t}$$

7. *See* [MARGOLIS 1995a] for a discussion of how three different types of compensation models (implicit contracts à la [BEAUDRY and DINARDO 1991] or [HARRIS and HOLMSTROM 1982], matching à la [JOVANOVIC 1979], and screening à la [SALOP and SALOP 1976] or [MARGOLIS 1995b]) can generate cohort effects in compensation policies and in returns to seniority.

In this specification, i denotes an individual, and T_0 is the date at which the individual joined her current employer. t measures years of employment with the current employer, and thus $s_{i,T_0+t} = t$ refers to the seniority of individual i hired at date T_0 and observed t years after being hired. The function $J(i, T_0 + t)$ identifies the firm j at which individual i works at date $T_0 + t$. In other words, if individual a was hired by firm A in 1980 and is observed in 1986 as still employed at firm A, then $i = a$, $T_0 = 1980$, $t = s_{a,1986} = 6$ and $J(a, 1986) = A$.

We denote the total annual compensation of individual i observed at the date $T_0 + t$ as w_{i,T_0+t}. X_{i,T_0+t} corresponds to a set of individual-specific characteristics that are observable to both the econometrician and the market and $X_{i,T_0+t}\beta$ is the market value of these characteristics. α_i is an individual fixed effect, representing the market value of individual-specific characteristics that are observed by the market but unobserved by the econometrician. If the econometrician were to observe all of the characteristics that were valued by the market, $\alpha_i = \alpha_j = 0$ for all i, j. In models with identical individual effects, $\alpha_i = \alpha_j = \alpha$ for all i, j, and α becomes part of the intercept in the X_{i,T_0+t} matrix. For this reason, one typically imposes the identification restriction that $\sum_i \alpha_i = 0$ in models with varying individual effects.

$\phi_{J(i,T_0+t),T_0}$ is the cohort-enterprise specific fixed effect (intercept) and $\gamma_{J(i,T_0+t),T_0}$ is the cohort-enterprise specific linear return to job seniority t [8]. Note that in a model with identical firms, $\phi_{j_1,T_0} = \phi_{j_2,T_0} = \phi_{T_0}$ and $\gamma_{j_1,T_0} = \gamma_{j_2,T_0} = \gamma_{T_0}$ for all j_1, j_2. In a model without cohort effects $\phi_{j,T_0} = \phi_{j,T_1} = \phi_j$ and $\gamma_{j,T_0} = \gamma_{j,T_1} = \gamma_j$ for all T_0, T_1. In a model with neither cohort effects nor heterogeneous firms, $\phi_{j_1,T_0} = \phi_{j_2,T_1} = \phi$ and $\gamma_{j_1,T_0} = \gamma_{j_2,T_1} = \gamma$ for all j_1, j_2 and for all T_0, T_1. In this last model, ϕ is once again subsumed into the intercept in X_{i,T_0+t}.

3 The Data

The data used in the empirical work below are the same as those described in ABOWD, KRAMARZ and MARGOLIS [1994] [9]. These data, derived from the DAS (Déclarations annuelles de salaire, or Annual Salary Reports)

8. In the models we estimate below, we typically consider specifications with both a linear return to seniority (as in equation 1.1) and with a return to seniority that includes higher-order terms as well. MURPHY and WELCH [1992] suggest that a specification with terms above the second order is appropriate, and where possible we will present specification tests that support this conclusion.

9. ABOWD, KRAMARZ and MARGOLIS [1994] provide a lengthy data appendix describing the construction of the data set in detail. Thus we will only briefly describe the construction of the data set here.

are a panel of employer-employee matched observations with identifying information sufficient to follow both firms and individuals over time, collected by INSEE (the French National Statistics Institute). The data cover the period 1976-1987, although observations from 1981 and 1983 were not made available by INSEE. They cover all workers employed in France born in October of an even-numbered year, thus ensuring a random initial sample of individuals. The data made available concern all such people working in enterprises with 10 employees or more, and French government employees are also excluded (although employees of government-owned enterprises are present in the data).

These data were cleaned [10] and compiled into a database containing initially 5,325,352 observations, of which 64.7 percent were from men. A simple regression of the form

$$w_{i,T_0+t} = X_{i,T_0+t}\beta + \varepsilon_{i,T_0+t}$$

was run on these data, where w_{i,T_0+t} represents the log of the real annualized total compensation cost (measured in thousands of 1980 french francs), and the vector X_{i,T_0+t} includes the variables male, Paris region, experience through experience to the fourth power [11], 7 education-level indicators, and 9 year indicators. All observations more than 5 standard deviations away from their predicted values were considered outliers. These observations were eliminated from the data set on which the more complicated estimators (those of TOPEL [1991], ABOWD, KRAMARZ and MARGOLIS [1994] and the cohort-based estimates) were calculated, although both data sets were used in the OLS and individual fixed effect regressions. Table 1 provides some basic sample statistics for both of the data sets used.

Looking at table 1, one can see that the outliers only comprised 0.38 percent of the original sample. The outliers-eliminated sample is extremely close to the overall sample, with most of the outliers apparently being unreasonably low earners as opposed to unreasonably high earners. The eliminated outliers also seem to be slightly more experienced and have slightly lower job seniority on average, but one cannot reject the hypothesis that the two samples are drawn from the same population with any generally accepted degree of confidence. In the OLS and individual fixed effect

10. *See* the data appendix of ABOWD, KRAMARZ and MARGOLIS [1994] for details on the preliminary treatment of the data.

11. Experience is measured as the number of years since the end of schooling (prior to 1976) plus the cumulated actual labor force experience (from 1976 onwards). *See* the data appendix of ABOWD, KRAMARZ and MARGOLIS [1994] for details.

TABLE 1

Descriptive Statistics

Variable	All Observations Mean	All Observations Std. Dev.	Outliers Eliminated Mean	Outliers Eliminated Std. Dev.
Log(Total Compensation Cost)	4.251	0.556	4.258	0.519
Seniority	6.379	7.138	6.391	7.140
Seniority ^{2}/100	0.916	1.485	0.918	1.486
Seniority ^{3}/1000	1.590	3.290	1.593	3.292
Seniority ^{4}/10000	3.044	7.721	3.050	7.727
Experience	16.610	11.927	16.610	11.923
Experience ^{2}/100	4.181	4.938	4.180	4.935
Experience ^{3}/1000	12.615	19.552	12.608	19.535
Experience ^{4}/10000	41.869	78.919	41.834	78.825
Elementary School	0.168	0.157	0.168	0.157
Junior High School	0.067	0.089	0.067	0.089
High School Grad	0.059	0.084	0.059	0.084
Basic Vo-Tech	0.239	0.178	0.240	0.178
Advanced Vo-Tech	0.064	0.087	0.064	0.087
Tech U/Undergrad	0.060	0.097	0.060	0.097
Grad School	0.039	0.085	0.039	0.085
1977	0.100	0.300	0.100	0.300
1978	0.102	0.303	0.103	0.303
1979	0.104	0.305	0.104	0.305
1980	0.104	0.305	0.104	0.305
1982	0.102	0.302	0.102	0.302
1984	0.098	0.297	0.098	0.297
1985	0.097	0.295	0.097	0.295
1986	0.098	0.298	0.098	0.298
1987	0.099	0.299	0.099	0.299
Male	0.647	0.478	0.647	0.478
Paris Region	0.269	0.443	0.268	0.443
	n=5,325,352		n=5,305,108	

Notes: The outliers eliminated sample of the DAS eliminated all observations more than 5 standard deviations away from the predicted wage in an OLS wage regression of log real annualized total compensation cost on sex, Paris region, experience-experience ^{4}, 7 education indicators and 9 year indicators. The French education levels were translated into their closest U.S. equivalents.

models, we present Hausman-Wu style tests of the equivalence of the coefficients of level seniority as a check on the representativeness of the outliers-eliminated sample [12].

12. We present results for both the full sample and the outliers eliminated sample in sections 4 and 5 for two reasons. First, we want to demonstrate that the two samples generate similar results for simple estimators of returns to job tenure. Second, although the full sample has more observations, ABOWD, KRAMARZ and MARGOLIS [1994] used the outliers-eliminated sample for their estimates. We would like to draw conclusions from our TOPEL [1991] 2-step and cohort-based estimators from the same data set, and thus we use the outliers-eliminated sample in sections 6, 7 and 8.

4 Ordinary Least Squares Estimates of Returns to Job Seniority

Our first set of estimates of returns to job seniority use a simple specification of the form

$$\log(w_{i,T_0+t}) = \beta X_{i,T_0+t} + \gamma s_{i,T_0+t} + \varepsilon_{i,T_0+t}. \tag{2}$$

There are three different assumptions that one might make to justify the elimination of α_i from the statistical model presented in equation (1), as we do here. First one might assume that the econometrician observes all of the relevant across-worker heterogeneity. Second, one might assume that all workers are identical in their cross-worker unobserved heterogeneity, in which case α is estimated as part of the model intercept. Third, one might suppose that the existing across-worker heterogeneity in wages is uncorrelated with job seniority and the elements of X_{i,T_0+t}. In this case we can consistently estimate β and γ without including the α_i term, and such a specification would have the mean value of the unobserved and uncorrelated individual specific heterogeneity being subsumed into the model intercept.

Next, note that there are no j-specific subscripts anywhere. We are thereby assuming that either there is no across-firm heterogeneity in intercepts ϕ or seniority returns γ, or that any across-firm heterogeneity in ϕ is uncorrelated with the X_{i,T_0+t} and the economy-wide γ. Second, there are no cohort-specific (T_0) terms anywhere. Thus if there was no heterogeneity in ϕ_j, the across-firm new-employer intercept term ϕ would be subsumed into the intercept in the regression. If we assumed uncorrelated heterogeneity, the across-firm mean ϕ_j would be a part of the intercept. The assumption that there are no cohort-specific seniority returns, coupled with the assumption that there is no heterogeneity across firms in seniority returns, means that one can estimate a single return to seniority γ for the entire economy.

We estimate two different specifications below. In model 1 we include only level seniority, whereas in model 2 we estimate returns to seniority through seniority to the fourth power (rescaled). Other variables included in the matrix X_{i,T_0+t} were total labor market experience through experience to the fourth power (rescaled), seven indicators for educational attainment (highest degree terminated), 9 year indicators, an indicator variable for a job in the Paris metropolitan area, and an indicator for the sex of the individual. The dependent variable was measured as the log of real annualized total compensation cost (w). Each model was run on two data sets, the first containing all observations and the second being identical to the first with the outliers eliminated. Table 2 presents the results of each of the two OLS regressions on each of the two data sets.

A Hausman-Wu style test of equality between the model 1 level seniority coefficient on the full and outliers-eliminated data sets cannot be rejected, nor can the equivalent test on the level seniority coefficient in model 2. This suggests that, as suspected, the relation between level seniority and

TABLE 2

OLS Results (Standard Deviations in Parentheses)

Variable	All Observations Model 1	All Observations Model 2	Outliers Eliminated Model 1	Outliers Eliminated Model 2
Intercept	3.271 (1.25E-03)	3.265 (1.26E-03)	3.287 (1.14E-03)	3.284 (1.15E-03)
Seniority	0.014 (3.50E-05)	0.069 (3.22E-04)	0.013 (3.18E-05)	0.061 (2.93E-04)
Seniority ^{2}/100	0 (0)	-0.931 (6.02E-03)	0 (0)	-0.817 (5.47E-03)
Seniority ^{3}/1000	0 (0)	0.494 (3.75E-03)	0 (0)	0.436 (3.41E-03)
Seniority ^{4}/10000	0 (0)	-0.081 (7.36E-04)	0 (0)	-0.072 (6.68E-04)
Experience	0.046 (2.81E-04)	0.034 (2.90E-04)	0.046 (2.55E-04)	0.035 (2.64E-04)
Experience ^{2}/100	-0.184 (2.56E-03)	-0.072 (2.63E-03)	-0.180 (2.33E-03)	-0.083 (2.39E-03)
Experience ^{3}/1000	0.038 (8.66E-04)	0.003 (8.85E-04)	0.038 (7.87E-04)	0.006 (8.05E-04)
Experience ^{4}/10000	-0.003 (9.64E-05)	0.000 (9.83E-05)	-0.003 (8.77E-05	0.000 (8.95E-05)
Elementary School	0.040 (9.18E-04)	0.046 (1.72E-03)	0.039 (8.34E-04)	0.048 (1.57E-03)
Junior High School	0.084 (9.14E-04)	0.391 (2.44E-03)	0.077 (8.30E-04)	0.394 (2.22E-03)
High School Grad	0.108 (9.13E-04)	0.587 (2.80E-03)	0.100 (8.29E-04)	0.596 (2.55E-03)
Basic Vo-Tech	0.113 (9.13E-04)	0.226 (1.45E-03)	0.107 (8.29E-04)	0.225 (1.32E-03)
Advanced Vo-Tech	0.140 (9.21E-04)	0.577 (2.48E-03)	0.134 (8.36E-04)	0.579 (2.25E-03)
Tech U/Undergrad	0.178 (9.33E-04)	0.580 (2.39E-03)	0.170 (8.48E-04)	0.582 (2.17E-03)
Grad School	0.199 (9.37E-04)	1.351 (2.75E-03)	0.190 (8.51E-04)	1.373 (2.50E-03)
1977	0.214 (9.34E-04)	0.059 (9.26E-04)	0.207 (8.49E-04)	0.054 (8.41E-04)
1978	0.221 (9.34E-04)	0.103 (9.27E-04)	0.216 (8.49E-04)	0.093 (8.42E-04)
1979	0.061 (1.73E-03)	0.125 (9.28E-04)	0.062 (1.57E-03)	0.114 (8.43E-04)
1980	0.401 (2.45E-03)	0.127 (9.31E-04)	0.404 (2.22E-03)	0.117 (8.45E-04)
1982	0.599 (2.81E-03)	0.148 (9.41E-04)	0.608 (2.55E-03)	0.138 (8.55E-04)
1984	0.246 (1.45E-03)	0.183 (9.55E-04)	0.244 (1.31E-03)	0.172 (8.67E-04)
1985	0.595 (2.48E-03)	0.202 (9.59E-04)	0.596 (2.25E-03)	0.190 (8.71E-04)
1986	0.586 (2.40E-03)	0.216 (9.57E-04)	0.589 (2.18E-03)	0.206 (8.70E-04)
1987	1.362 (2.75E-03)	0.224 (9.58E-04)	1.384 (2.51E-03)	0.215 (8.71E-04)
Male	0.206 (4.60E-04)	0.204 (4.61E-04)	0.203 (4.18E-04)	0.201 (4.19E-04)
Paris Region	0.139 (4.74E-04)	0.142 (4.73E-04)	0.142 (4.31E-04)	0.145 (4.30E-04)
	R^2=0.293	R^2=0.297	R^2=0.331	R^2=0.337
	n=5,325,352		n=5,305,108	

H_0: γ(Model 1, All Observations) = γ(Model 1, Outliers Eliminated)	$\chi^2(1)$=0.29
H_0: γ(Model 2, All Observations) = γ(Model 2, Outliers Eliminated)	$\chi^2(1)$=2.50
H_0: γ(Model 1, All Observations) = γ(Model 2, All Observations)	$\chi^2(1)$=11.56
H_0: γ(Model 1, Outliers Eliminated) = γ(Model 2, Outliers Eliminated)	$\chi^2(1)$=9.55

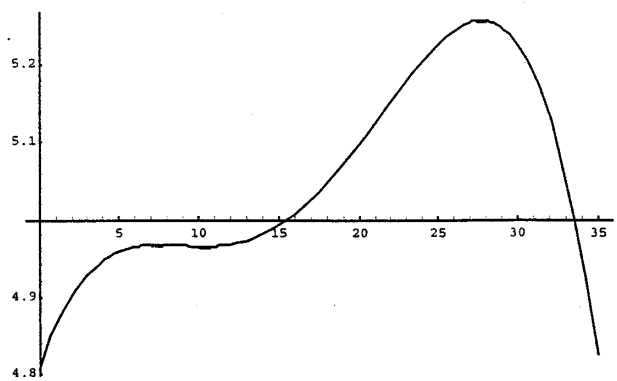
Notes: The outliers eliminated sample of the DAS eliminates all observations more than 5 standard deviations away from the predicted wage in an OLS wage regression of log real annualized total compensation cost on sex, Paris region, experience-experience ^{4}, 7 education indicators and 9 year indicators. The French education levels were translated into their closest U.S. equivalents.

earnings is not affected by the elimination of outliers. On the other hand, Hausman-Wu tests for equality of the model 1 and model 2 coefficients on level job seniority are rejected for both the full and outliers-eliminated samples. This suggests that the most simple OLS model, with only level seniority, is misspecified.

Figure 1 shows the how predicted earnings vary with seniority for a male high-school graduate living in Paris with 35 years of work experience

FIGURE 1

OLS Returns to Job Seniority



observed in 1980. It is interesting to note that, among workers with 35 years of work experience, it is not those that have been employed the longest that would have the highest predicted earnings, but rather who obtained their job with their current employer after having been on the job market for about 8 years. We will discuss the implications for such a profile, if it is robust, below.

5 Individual Fixed Effects Estimates of Returns to Job Seniority

Of course, the assumptions made to justify the OLS specification are extremely restrictive. In particular, in the OLS model it is likely that $E(s'\varepsilon) \neq 0$ due to the presence of omitted variables that might be correlated with job seniority, and thus the OLS estimate of returns to seniority is likely to suffer from substantial omitted variable bias. One source of such correlation, for example, could be the possibility that more able workers are the last ones fired in the event of a negative shock to the firm. If true, this generates a relation between seniority and earning via the omitted variable ability, and implied an upward bias in the naive OLS estimate of returns to job seniority.

As a first step toward resolving the problem, we next estimate returns to job seniority using a specification that allows for individual specific heterogeneity to be correlated with seniority and elements of the X_{i,T_0+t} matrix. The model we estimate is

$$\log(w_{i,T_0+t}) = \beta X_{i,T_0+t} + \alpha_i + \gamma s_{i,T_0+t} + \varepsilon_{i,T_0+t}.$$

Here we are once again making the same sort of assumptions concerning the ϕ_j and γ_j terms as in section 4 above. However, now we are explicitly allowing the individual specific unobserved heterogeneity to be correlated with the other observed characteristics, including seniority. In order to control for the individual specific effects, we project the X_{i,T_0+t}, t and $\log(w_{i,T_0+t})$ onto their individual-specific means. We thus estimate an equation of the form

$$\log(\tilde{w}_{i,T_0+t}) = \beta\tilde{X}_{i,T_0+t} + \gamma\tilde{s}_{i,T_0+t} + \tilde{\varepsilon}_{i,T_0+t} \tag{3}$$

where

$$\begin{aligned}
\log(\tilde{w}_{i,T_0+t}) &= \log(w_{i,T_0+t}) - \log(\bar{w}_i) \\
\tilde{X}_{i,T_0+t} &= X_{i,T_0+t} - \bar{X}_i \\
\tilde{s}_{i,T_0+t} &= s_{i,T_0+t} - \bar{s}_i \\
\tilde{\varepsilon}_{i,T_0+t} &= \varepsilon_{i,T_0+t} - \bar{\varepsilon}_i
\end{aligned}$$

and the terms $\log(\bar{w}_i)$, $\bar{X}_i$, $\bar{s}_i$ and $\bar{\varepsilon}_i$ refer to means of the relevant variables over all observations corresponding to individual i [13]. Under our assumptions made above, OLS run on this specification generates unbiased consistent estimates of β and γ [14]. We use the same data as used in estimating equation (2). The results of the estimations of each of the two OLS regressions on each of the two data sets are found in table 3.

As suggested above, the estimated returns to seniority that one finds in the OLS estimates (coefficients on level seniority of 0.069 on the whole sample, 0.061 for the sample with outliers eliminated in the preferred specifications) are greatly reduced when one includes individual specific effects (coefficients on level seniority of 0.032 on the whole sample and 0.029 on the outliers-eliminated sample in the preferred specifications). Hausman-Wu tests reject the OLS specification for both the full sample and for the outliers eliminated sample in the preferred specifications. The apparent positive bias in the estimates of returns to seniority resulting from the omission of individual specific effects is important. As mentioned above, it suggests that workers that have unobserved individual specific characteristics that are worth more on the job market are likely to also be the workers with the highest job seniority. This makes sense if α_i reflects additional unmeasured productive capacity and there are, say, positive costs to employer search or screening of workers. In this case it is costly to find good (high-α_i) workers, and thus these workers will be last ones laid off in the event of a negative shock to the firm. The possibility that the workers are paid a fixed share of their marginal productivity independent of employer, and in particular with their current employer, means that they will be no more nor less likely to separate for employee-induced reasons. Thus the negative correlation between α_i and hazard rates will induce a

13. We are assuming that ε_{i,T_0+t} is distributed such that not only is $E(\varepsilon_{i,T_0+t}) = 0$, but also $E(\varepsilon_{I,T_0+t} \mid I = i) = 0$ for all i.
14. *See* HSIAO [1986] for a proof.

TABLE 3

Individual Fixed Effects Results (Standard Deviations in Parentheses)

Variable	All Observations Model 1	All Observations Model 2	Outliers Eliminated Model 1	Outliers Eliminated Model 2
Seniority	0.003 (4.58E-05)	0.032 (2.86E-04)	0.003 (3.93E-05)	0.029 (2.45E-04)
Seniority ^{2}/100	0 (0)	-0.467 (5.28E-03)	0 (0)	-0.425 (4.53E-03)
Seniority ^{3}/1000	0 (0)	0.247 (3.24E-03)	0 (0)	0.227 (2.78E-03)
Seniority ^{4}/10000	0 (0)	-0.042 (6.27E-04)	0 (0)	-0.039 (5.37E-04)
Experience	0.055 (3.84E-04)	0.046 (3.97E-04)	0.054 (3.31E-04)	0.046 (3.41E-04)
Experience ^{2}/100	-0.347 (2.75E-03)	-0.274 (2.86E-03)	-0.343 (2.37E-03)	-0.278 (2.46E-03)
Experience ^{3}/1000	0.080 (9.69E-04)	0.056 (1.00E-03)	0.079 (8.32E-04)	0.057 (8.62E-04)
Experience ^{4}/10000	-0.007 (1.12E-04)	-0.004 (1.15E-04)	-0.007 (9.64E-05)	-0.005 (9.89E-05)
1977	0.023 (6.96E-04)	0.035 (7.08E-04)	0.024 (5.98E-04)	0.035 (6.08E-04)
1978	0.063 (7.90E-04)	0.079 (8.09E-04)	0.063 (6.79E-04)	0.077 (6.95E-04)
1979	0.093 (9.24E-04)	0.110 (9.45E-04)	0.092 (7.94E-04)	0.107 (8.12E-04)
1980	0.107 (1.08E-03)	0.124 (1.10E-03)	0.108 (9.28E-04)	0.123 (9.47E-04)
1982	0.151 (1.43E-03)	0.169 (1.45E-03)	0.151 (1.23E-03)	0.167 (1.25E-03)
1984	0.200 (1.79E-03)	0.220 (1.81E-03)	0.199 (1.54E-03)	0.217 (1.56E-03)
1985	0.224 (1.97E-03)	0.244 (1.98E-03)	0.222 (1.69E-03)	0.240 (1.71E-03)
1986	0.256 (2.15E-03)	0.277 (2.17E-03)	0.256 (1.85E-03)	0.274 (1.87E-03)
1987	0.277 (2.34E-03)	0.298 (2.36E-03)	0.278 (2.02E-03)	0.297 (2.03E-03)
Paris Region	0.072 (1.08E-03)	0.073 (1.08E-03)	0.081 (9.31E-04)	0.082 (9.29E-04)
	R^2=0.741	R^2=0.742	R^2=0.782	R^2=0.783
	n=5,325,352		n=5,305,108	

H_0: γ(Model 2, All Observations, OLS) = γ(Model 2, All Observations, Fixed Effects)	$\chi^2(1)$=39.33
H_0: γ(Model 2, Outliers Eliminated, OLS) = γ(Model 2, Outliers Eliminated, Fixed Effects)	$\chi^2(1)$=21.65

Notes: The outliers eliminated sample of the DAS eliminates all observations more than 5 standard deviations away from the predicted wage in an OLS wage regression of log real annualized total compensation cost on sex, Paris region, experience-experience ^{4}, 7 education indicators and 9 year indicators. The French education levels were translated into their closest U.S. equivalents.

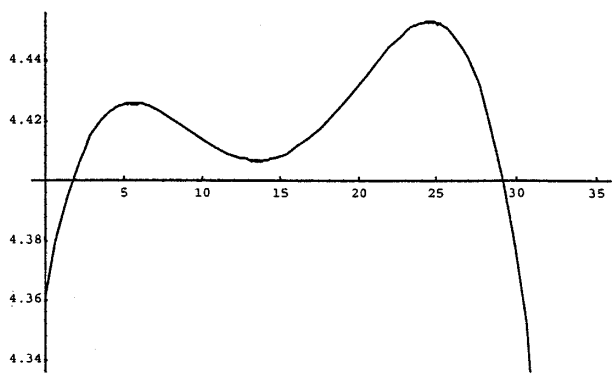
positive correlation between α_i and job seniority, and thereby a positive bias in the naive OLS estimates.

Figure 2 shows the equivalent picture to figure 1, but this time using the coefficients from the individual fixed effects estimates in table 3. Note how, as in the case of OLS, the highest earnings are attained at some level of seniority less than the maximum of 35 years. In this case the largest earnings seem to be had by workers with 25 years of experience, and there is a second local maximum at around 6 years of experience. The fact that earnings seem to decrease with seniority after 5 years until about 15 years is also somewhat unusual, although the OLS results imply a marginal dip in earnings to a local minimum at about 11 years of seniority. As far as we can tell, no single theory of compensation determination predicts such a double-hump phenomenon.

Note, however, that neither of these estimators makes any allowance for firm specific heterogeneity in compensation. This omission could cause both

FIGURE 2

Individual Fixed Effects Returns to Job Seniority



the OLS and individual fixed effects estimators of returns to job seniority to suffer from omitted variable bias [15]. In the following three sections we estimate returns to seniority, allowing for firm specific heterogeneity of increasing complexity, using the multiple step procedures suggested by TOPEL [1991] and ABOWD, KRAMARZ and MARGOLIS [1994], and we also estimate the firm specific heterogeneity parameters ϕ_j and γ_j separately for different cohorts within a firm.

6 The TOPEL [1991] Two-Step Estimator

We next estimate returns to job seniority using the method suggested in TOPEL [1991]. In this approach, one first estimates a combined return to job seniority and total labor market experience using within-employer differenced data. One then decomposes the combined estimate into components that arise due to job seniority and components coming from experience. Using the notation of equation (1), Topel's procedure generates an unbiased estimate of the combined return $(\gamma + \beta)$ under the assumption that $\gamma_j = \gamma_k = \gamma$ for all j, k. He also suggests a way to approximate the sign and size of one component of the bias in estimated wage growth.

15. *See* TOPEL [1991] for a detailed discussion of how ignoring heterogeneity in ϕ_j can bias estimated returns to seniority. MARGOLIS [1995b] also provides a model in which the heterogeneity in returns to seniority γ_j and in firm specific intercepts ϕ_j is intimately linked to differences in average seniority and in compensation.

TABLE 4

Descriptive Statistics for Differenced Data

Variable Name	Differenced Data Mean	Std. Dev.
Δw	0.035	0.270
Seniority, Experience	0.963	0.149
Seniority ^{2}/100	0.164	0.143
Seniority ^{3}/1000	0.355	0.471
Seniority ^{4}/10000	0.815	1.411
Experience ^{2}/100	0.361	0.233
Experience ^{3}/1000	1.382	1.488
Experience ^{4}/10000	5.556	7.897
	n=2,517,026	

Notes: These data are drawn from the outliers eliminated subsample of the DAS. Differences are calculated within individual-firm matches across time.

In the first step of the estimation process, we calculated first differences within individual-firm pairs for the outliers-eliminated data. Sample statistics for this data set appear in table 4 [16]. We then regressed the first differenced log real total annualized compensation cost on the first differenced seniority through first differenced seniority to the fourth power and first differenced total labor market experience through first differenced experience to the fourth power. This model was estimated in two forms: with the parameter restriction $\gamma_2 = \gamma_3 = \gamma_4 = 0$ (only level seniority included, denoted Model T1A below) and without any parameter restrictions (Model T1B below) [17]. The results of these first-step models are shown in table 5 below.

A Hausman-Wu test for equality of the intercept term across the two specifications is rejected at most levels of significance. Thus we once again tend to favor the more detailed specification (model T1B), but for the sake of comparability with the results in sections 4 and 5 and with the estimates to follow in sections 7 and 8, we present both sets of results.

Exploiting the identity $Exp_{i,T_0+t} = Exp_{i,T_0} + t$, TOPEL [1991] showed that one can separate the estimated coefficient on seniority into a component that can be attributed to seniority returns and a component that can be attributed to experience returns. This involves regressing the difference between current log compensation and predicted log compensation (according the results from the first step estimation) on total labor market experience at

16. Note that the coefficient on differenced seniority is not exactly 1. This is because the data allowed us to determine what fraction of the year and individual was employed with the same firm. Factors that allow continued employment with the same firm in which seniority and experience increase by less that 1 each year (temporary layoffs, for example) might be the source of the difference. *See* TOPEL [1991] for a discussion of the impact that this might have on estimated returns to seniority.

17. Note that, in this specification, $\gamma_1, \gamma_2, \gamma_3$ and γ_4 are coefficients shared by all firms, and do not represent returns to seniority for 4 different firms (as would be the case using the notation in equation (1)).

TABLE 5

First Step Topel Model Results (Std. Dev. in Parentheses).

Variable	Differenced Data Model T1A	Model T1B
Seniority, Experience	0.087 (6.83E-04)	0.106 (7.44E-04)
Seniority ^{2}/100	0 (0)	-0.551 (9.32E-03)
Seniority ^{3}/1000	0 (0)	0.300 (6.21E-03)
Seniority ^{4}/10000	0 (0)	-0.052 (1.25E-03)
Experience ^{2}/100	-0.297 (6.46E-03)	-0.231 (6.60E-03)
Experience ^{3}/1000	0.057 (2.28E-03)	0.040 (2.34E-03)
Experience ^{4}/10000	-0.004 (2.61E-04)	-0.002 (2.66E-04)
	R^2=0.022	R^2=0.020
H_0: $\gamma_1+\beta_1$(Model T1A) = $\gamma_1+\beta_1$(Model T1B)		$\chi^2(1)$=5.92

Notes: These data are drawn from the outliers eliminated subsample of the DAS. Differences are calculated within individual-firm matches across time.

the date the job was started (Exp_{i,T_0}) and a vector of other individual specific variables that might be correlated with earnings. We thus estimated the model

$$\log(\tilde{w}_{i,T_0+t}) = \beta_1 Exp_{i,T_0} + \theta F_{i,T_0+t} + e_{i,T_0+t}$$

where

$$\begin{aligned}\log(\tilde{w}_{i,T_0+t}) = \log(w_{i,T_0+t}) &- (\widehat{\gamma_1+\beta_1})t_{i,j,T_0+t} - \hat{\gamma}_2 t^2_{i,j,T_0+t} - \hat{\gamma}_3 t^3_{i,j,T_0+t} \\ &- \hat{\gamma}_4 t^4_{i,j,T_0+t} - \hat{\beta}_2 Exp^2_{i,T_0+t} - \hat{\beta}_3 Exp^3_{i,T_0+t} \\ &- \hat{\beta}_4 Exp^4_{i,T_0+t}\end{aligned}$$

and F_{i,T_0+t} corresponds to the other individual-specific observable components, namely the elements of X_{i,T_0+t} not previously included, such as education, sex and region. The results of estimations of this model based on the restricted (Model T2A below) and unrestricted (Model T2B below) first-step estimates appear in table 6 below.

TOPEL [1991] noted that although his estimator of joint returns to seniority and returns to experience initial experience $(\widehat{\gamma_1+\beta_1})$ would be consistent, endogenous job changing would generate a bias in estimated wage growth that could be evaluated. He also suggested a method for estimating the part of the bias in the estimate of γ_1 (and thereby in the estimate of β_1) due to covariance of job tenure with unobserved characteristics that determine ϕ_j, since his approach provides no asymptotically unbiased method for dividing the sum $(\widehat{\gamma_1+\beta_1})$ into γ_1 and β_1. Table 7 below uses Topel's preferred method to break down the estimates of the joint returns to seniority and experience (in levels shown in table 5) generated by each of the two models (Model A with the parameter restrictions and Model B unrestricted) into seniority and experience components, as well as estimates of the bias in

TABLE 6

Second Step Topel Model Results (Std. Dev. in Parentheses).

Variable	All Observations (Outliers Removed) Model T2A	Model T2B
Initial Experience	0.066 (3.42E-05)	0.061 (3.29E-05)
Elementary School	2.023 (2.51E-03)	2.062 (2.42E-03)
Junior High School	2.644 (3.86E-03)	2.682 (3.72E-03)
High School Grad	2.433 (4.57E-03)	2.475 (4.40E-03)
Basic Vo-Tech	2.263 (1.96E-03)	2.299 (1.88E-03)
Advanced Vo-Tech	1.738 (4.02E-03)	1.801 (3.87E-03)
Tech U/Undergrad	2.253 (3.81E-03)	2.329 (3.67E-03)
Grad School	1.822 (4.53E-03)	1.900 (4.36E-03)
1977	1.249 (1.40E-03)	1.245 (1.34E-03)
1978	1.374 (1.38E-03)	1.362 (1.33E-03)
1979	1.447 (1.37E-03)	1.429 (1.32E-03)
1980	1.504 (1.36E-03)	1.480 (1.31E-03)
1982	1.643 (1.37E-03)	1.608 (1.31E-03)
1984	1.790 (1.38E-03)	1.744 (1.32E-03)
1985	1.787 (1.38E-03)	1.742 (1.33E-03)
1986	1.802 (1.37E-03)	1.758 (1.32E-03)
1987	1.820 (1.37E-03)	1.776 (1.32E-03)
Male	0.484 (7.48E-04)	0.496 (7.20E-04)
Paris Region	0.296 (7.94E-04)	0.287 (7.65E-04)
	R^2=0.969	R^2=0.970
	n=5,305,108	

Notes: Initial experience is measured as total labor market experience in the year that the employee joined the firm. Model T2A refers to the second step of the linear specification in seniority, and model T2B refers to the second step of the quartic specification in seniority.

TABLE 7

Estimated Returns Using the Topel Two-Step Estimator.

Model Type	Total Returns	Seniority	Experience	Wage Growth Bias
Model A	0.087	0.021	0.066	-0.052
Model B	0.106	0.045	0.061	-0.040
TOPEL [1991]	0.126	0.054	0.071	0.002

Notes: Model A refers to the linear specification in seniority, and model B refers to the quartic specification in seniority. Wage growth bias refers to the estimate of the bias in cross section estimates of the sum of returns to seniority and experience.

cross-sectional estimates of wage growth (the sum of the biases on γ_1 and β_1), and compares them to the results in TOPEL [1991] from the United States.

The Hausman-Wu tests following the first step regression suggest that, consistent with the OLS and the fixed effects estimators, including the higher order terms in seniority has a dramatic effect on estimated returns to seniority. Comparing our results with those found in TOPEL [1991], we see

that returns to seniority in France seem to be lower than in the U.S., although the results for returns to total labor market experience are comparable.

Interpreting the ϕ_j term in equation (1) as a match specific component of earnings (i.e. ϕ_{ij}), TOPEL [1991] suggests potential signs for the biases in cross section estimates of returns to seniority and experience based on endogenous separations and firm-to-firm movements. Workers start their careers with a particular draw from the ϕ_{ij} distribution, and when a new offer ϕ_{ik} arrives such that $\phi_{ik} > \phi_{ij}$, the individual switches employers. Assuming that draws of ϕ_{ij} are i.i.d., the longer a person has been on the job market, the higher will be her wage, thus implying a positive correlation between ϕ_{ij} and total labor market experience.

The implications for the bias on job seniority are less clear. TOPEL [1991] suggests that if $\gamma > 0$, then individuals will need even higher ϕ_{ik} to be incited to switch employers. Thus the individuals with long tenure are those for which $\phi_{ik} - \phi_{ij}$ is not large enough to compensate for the loss in earnings due to "restarting the seniority clock". Conversely, those individuals with short tenure will be those who just received a big jump in ϕ_{ij}. These two effects imply a downward bias in estimated returns to seniority in a cross section (as in our section 4 and 5 estimators).

An alternative explanation is based exclusively on the increasing ϕ_{ij} over the career. The higher the current ϕ_{ij}, the less probable that a new draw of ϕ_{ij} will be sufficiently high to induce the worker to move. This implies that seniority will be short for matches with relatively low ϕ_{ij} and thus relatively low earnings. Later in the career, when the individual has sampled several ϕ_{ij} and finds herself employed at the firm that offered the highest ϕ_{ij}, tenure spells will be long and wages higher. Given that the ϕ_{ij} is specific to the individual-firm match, one can observe different individuals in the same firm with different expected seniorities, where the expected seniority will be positively correlated with ϕ_{ij} and experience will be also be positively correlated with ϕ_{ij}. This implies that the OLS and individual fixed effect estimates of returns to job seniority will be subject to upward biases.

Comparing the results in table 7 with those in table 3, one notes that the table 7 estimates are uniformly larger than the individual fixed effects estimates (0.045 versus 0.029 for level seniority, 0.061 vs 0.046 for level experience). Although the result on total experience seems inconsistent with endogenous switching models, the result on job seniority could be due to an explanation similar to Topel's. On the other hand, a regression of current tenure on initial experience yields a coefficient of 0.75 (contrary to -0.25 for the United States). This implies that, combined with the estimates of the wage growth bias in table 7, the estimate of γ_1 will be biased upwards by 0.039 in model A and 0.030 in model B due to covariance of job tenure with unobservables that determine the firm effect ϕ_j. This bias is large and positive, and could (in the case of model A) mask an underlying zero return to seniority. As noted in the second explanation above, however, this upward bias will be offset at least partially by a downward bias due to improving match quality over a career (this biases up the estimate of returns to total labor market experience and thus biases down the estimate of returns to seniority). Topel provides no insight on how to sign this bias term, although a negative value for the bias term suggests that true returns

to level experience are higher than estimated, and thus returns to seniority are lower than estimated.

The conclusion is unclear. Estimates of both returns to job seniority and returns to total labor market experience calculated according to the method suggested by TOPEL [1991] are above the estimates derived from fixed effects estimates. Acknowledging the fact that the partition of the joint returns is not consistently estimated, and since (under the equal γ and cohort-invariant ϕ and γ assumptions) the sum of the two returns is consistently estimated, this suggests that the negative bias in the fixed effect estimated returns to seniority outweighs the positive bias in the fixed effect returns to experience in the case of France. However, in the notation of equation (1), ϕ is firm specific and identical for all workers within the same firm. Thus Topel's and our explanations of the sources of bias in returns to tenure actually relate to the η_{i,T_0+t} in equation (1). A literal interpretation of his econometric procedure might impose the same constraints on ϕ as in the OLS and fixed effect estimators, while requiring us to treat equation (1) as having a match specific random effect in the "residual" η_{i,T_0+t}.

Lastly, the absence of any control for α_i in the second step estimates is another potential source of bias in the decomposition of the combined seniority-experience term. If workers with higher individual effects have longer seniorities (as suggested in section 5 above), then the share of joint returns attributed to seniority will be overestimated in the second step regression. The estimate of the joint return to job seniority and total labor market experience is still consistently estimated, since the first step regression used within-individual-firm differences. Thus although this suggests that the estimate of returns to seniority in the TOPEL [1991] method is biased upwards, since the sum of the two coefficients is unbiased, the wrong-signed bias on the returns to experience estimate remains unexplained.

7 The ABOWD, KRAMARZ and MARGOLIS [1994] Projection Method Estimator

All of these procedures have found significant (however small) returns to seniority in the fully saturated specifications on the outliers-eliminated data. They share a common assumption, however; namely that all firms reward seniority in the same manner ($\gamma_j = \gamma_k = \gamma$ for all j, k). If this assumption were to be violated, the estimated returns to seniority would be a sort of weighted average of the firm-specific returns. Thus even if only a few firms in the sample rewarded seniority and these firms contributed a high enough number of observations to the data set, we would run the risk of estimating positive economy-wide returns, and we might incorrectly assume that the majority of firms in the economy provided positive returns to job seniority.

ABOWD, KRAMARZ and MARGOLIS [1994] show that this is false [18]. Their estimation procedure allows them to account for individual- and firm-specific effects on earnings, as well as estimating firm-specific returns to seniority. Due to data limitations, however, they restrict their consideration to a specification equivalent to equation (1) with a linear spline in seniority (break point at 10 years of job seniority) under the assumption that $\phi_{j,T_0} = \phi_{j,T_1} = \phi_j$ and $\gamma_{j,T_0} = \gamma_{j,T_1} = \gamma_j$ for all T_0, T_1[19]. Using the same data as we use here (the outliers-eliminated sample), they find no significant returns to seniority on average (-3.37E-5 for men and 8.28E-4 for women across individuals). However, when looking across firms, they find a standard error in returns to seniority that is relatively large (0.077 relative to a mean across firms of 2.7E-3) and encompasses all of the above estimated returns within one standard deviation of the mean estimated returns to seniority. This implies that there is substantial heterogeneity in estimated returns to seniority, and thus the assumption inherent in the TOPEL [1991] procedure that $\gamma_j = \gamma_k = \gamma$ for all j, k needs to be reconsidered. Although it provides some useful information that different theoretical models can interpret, even the ABOWD, KRAMARZ and MARGOLIS [1994] projection method estimator ignores cohort effects. These cohort-firm specific returns to job seniority, and in particular their variance within firm or within cohort, provide information that can be of use for interpreting various models of compensation.

8 Cohort-Based Estimates of Returns to Job Seniority

With the knowledge that there is substantial across-firm heterogeneity in estimated returns to job seniority, we took the ABOWD, KRAMARZ and MARGOLIS [1994] projection method as a starting point for estimating directly the model described in equation (1). Because estimating cohort-firm specific intercepts ($\hat{\phi}_{j,T_0}$) and seniority returns ($\hat{\gamma}_{j,T_0}$) is particularly demanding on the data, we were not able to obtain large enough sample sizes for every cohort in every firm in the data. We were able to use 3,677,633

18. Since ABOWD, KRAMARZ and MARGOLIS [1994] used the same data we use here and estimate a model very similar to equation (1) - the only exception being that they also estimate a linear spline with a break at 10 years of seniority - we do not recalculate their results. What follows is a summary of the results that can be found in their paper. We use equation (1) and the previous discussion to interpret their results.

19. Because sample sizes necessary to estimate a full quartic specification for each firm are larger than those needed to estimate a simple linear spline, ABOWD, KRAMARZ and MARGOLIS [1994] chose the option of being able to estimate firm specific returns to seniority for a larger number of firms rather than that of estimating a full quartic specification for the fewer firms with enough data to meet the conditions necessary for identification.

of our 5,305,108 observations to calculate cohort-firm specific intercepts and slopes, 4,654,064 observations for the estimation of firm specific (but pooled across cohorts) slopes and intercepts (as in ABOWD, KRAMARZ and MARGOLIS [1994]), and all 5,305,108 observations for calculating cohort-specific (but pooled across firms) slopes and intercepts [20]. This distribution of observations across groups is encouraging in two respects. First, for 88 percent of the observations we are able to estimate at least a firm specific $\hat{\phi}_j$ and $\hat{\gamma}_j$, and 79 percent of these observations can be used to recover the relevant cohort-firm specific $\hat{\phi}_{j,T_0}$ and $\hat{\gamma}_{j,T_0}$ terms. Second, we can use *all* of the observations in estimating cohort specific intercepts and slopes. This is not surprising, given that the data set eliminated all individuals younger than 15 years old and older than 65 years old.

The ABOWD, KRAMARZ and MARGOLIS [1994] projection method was followed up to the point where they estimate firm-specific intercepts and seniority splines. Thus the estimated effects of education, total labor market experience, sex and region on the log of total real annual compensation found in ABOWD, KRAMARZ and MARGOLIS [1994] are unchanged. We then estimated equation (1) separately for each entering cohort in each firm in the cohort-firm group, and we pooled the remaining observations to estimate equation (1) for each cohort in the pool of firms. We also estimated equation (1) for each firm in the firm-only group [21]. For the cohort-only group, we estimated equation (1) by pooling all observations within the same cohort, independent of firm, and running the regression cohort by cohort.

Using the cohort-firm estimates, we then calculated the mean asymptotic within-firm and within-cohort variance in $\hat{\phi}_{j,T_0}$ and $\hat{\gamma}_{j,T_0}$. Asymptotic within variances were calculated as

$$\begin{aligned} var\left(\hat{\phi}_j\right) &= var_j\left(\sqrt{n_{j,T_0}}\,\hat{\phi}_{j,T_0}\right) \\ var(\hat{\gamma}_j) &= var_j\left(\sqrt{n_{j,T_0}}\,\hat{\gamma}_{j,T_0}\right) \\ var\left(\hat{\phi}_{T_0}\right) &= var_{T_0}\left(\sqrt{n_{j,T_0}}\,\hat{\phi}_{j,T_0}\right) \\ var(\hat{\gamma}_{T_0}) &= var_{T_0}\left(\sqrt{n_{j,T_0}}\,\hat{\gamma}_{j,T_0}\right) \end{aligned}$$

where var_j refers to the variance over all estimates in firm j, var_{T_0} refers to the variance over all estimates in entry cohort T_0, and n_{j,T_0} refers to the number of observations used to calculate the corresponding $\hat{\phi}_{j,T_0}$ or $\hat{\gamma}_{j,T_0}$. Table 8 presents these results. We can see that the average variance in $\hat{\phi}_{j,T_0}$ is much larger within cohorts than within firms (16.905 versus 1.058). On the other hand, the average variance in $\hat{\gamma}_{j,T_0}$ is not very different within

20. Clearly, the data used to estimate firm-cohort specific or firm specific ϕ and γ are more likely to come from observations of individuals employed in larger firms. Unfortunately there is no way to recover cohort-firm specific $\hat{\phi}_{j,T_0}$ and $\hat{\gamma}_{j,T_0}$ when there is not enough data present. Thus the estimation strategy implicitly imposes a sort of "ocean of small, identical firms" constraint, in which a single ϕ or γ is estimated for all observations coming from insufficiently present employers. This is equivalent to supposing that small firms have the same compensation policies and reward seniority identically.
21. This is essentially what is done in ABOWD, KRAMARZ and MARGOLIS [1994].

cohorts as opposed to within firms (1.380 versus 1.229). We also note that the negative correlation between $\hat{\phi}_j$ and $\hat{\gamma}_j$ observed by ABOWD, KRAMARZ and MARGOLIS [1994] is present even at the cohort-firm level, although the negative correlation is stronger across firms within a cohort (-0.860) than across cohorts within a firm (-0.724). This stronger relation within cohort than within firm is likely due to the larger variance of $\hat{\phi}_{j,T_0}$ within cohort than within firm.

TABLE 8

Analysis of the Cohort-Firm Estimates of ϕ and γ.

Criterion	Within Firm	Within Cohort
Average Variance in ϕ	1.058	16.905
Average Variance in γ	1.229	1.380
Average Correlation Coefficient (ϕ, γ)	-0.724	-0.860

Notes: The variance measures are average asymptotic variances within the particular firm or cohort. The within firm averages are calculated over 39,876 firms, while the within cohort averages are calculated over 43 cohorts.

9 Interpretation of the Empirical Results

The estimators presented in sections 4, 5 and 6 show that methods that impose identical returns to seniority across different firms and cohorts tend to predict positive returns, whereas when each firm and/or cohort is allowed its own compensation policy and return to seniority, the average return becomes zero. The projection based estimator of ABOWD, KRAMARZ and MARGOLIS [1994] is a good step in the right direction, but when we estimate slopes and intercepts separately for cohort-firms, we find still more variance.

The pattern of this variance, and in particular the characteristics of the negative covariances that we find, are quite interesting from a theoretical point of view. For example, consider the framework of compensation determination that is suggested by BEAUDRY and DINARDO [1991]. In this framework of implicit contracts with costless mobility, one ought to observe significant variance within firm by cohort in both the intercept $\hat{\phi}_{j,T_0}$ and the slope $\hat{\gamma}_{j,T_0}$. We only find moderate variance within firm in both intercepts and slopes (table 8). However, this theory suggests that all individuals of the same entry cohort where hired under the same general labor market conditions, and their wage growth is driven by the same general shocks. Thus there should be no variance within cohort across firms in slopes or intercepts. This implication is not borne out by the data for the $\hat{\phi}_{j,T_0}$ terms, although it seems more plausible for $\hat{\gamma}_{j,T_0}$. If one would like to relax this

theory, say by permitting an additional firm-specific shock and adding small mobility costs, one might be able to generate variance across firms within cohorts in the $\hat{\phi}_{j,T_0}$ terms, but there is no reason why this should only affect intercepts and not slopes. Thus the theory of implicit contracts with costless mobility, which could have rationalized the results of ABOWD, KRAMARZ and MARGOLIS [1994] by insisting on the fact that no controls were made for differences across firms in the cohort-composition of their workforces, has a harder time explaining the data when such controls are made and residual heterogeneity across firms within cohorts in compensation policies persists.

Another theoretical model, that of MARGOLIS [1995b], suggests that screening considerations could cause firms to offer different slopes and intercepts. The combination of endogenously chosen compensation policies and seniority returns plus variations in the quality of the entry cohorts will generate both heterogeneity across firms within cohorts (since different offered compensation policies and seniority returns attract different sorts of workers), and across cohorts within firms (since the initial composition of the cohort will be a function of supply-side conditions as well). This theory can accommodate our heterogeneity results for $\hat{\phi}_{j,T_0}$ within cohort with ease, and can explain the small average within firm variance in $\hat{\phi}_{j,T_0}$ by small variance in supply-side factors. Unfortunately, the large average variance in $\hat{\phi}_{j,T_0}$ within cohorts should be associated with a larger average variance in $\hat{\gamma}_{j,T_0}$ within cohorts, and this is less clear in the results presented in section 8. Similar experiments can be tried with other models that generate cohort effects [22], but to the best of our knowledge there is no single theory that can adequately explain the empirical results found in section 8.

10 Conclusion

This paper has demonstrated the implications of various simplifying assumptions on estimates of returns to job seniority, and has highlighted certain problems that were previously unknown by making use of variations in compensation policy and returns to seniority within firms across entering cohorts. A ranking of the order of magnitude of our results gives the highest estimates for OLS (6.9 percent increase per year), followed by the TOPEL [1991] two step estimator (4.5 percent per year), the individual fixed effects estimator (3.2 percent per year) and finally the ABOWD, KRAMARZ and MARGOLIS [1994] projection based estimator and a variant that estimates cohort-firm effects (both with essentially zero mean returns to seniority across firms). The remarkable heterogeneity in firm specific, and particularly cohort-firm specific compensation policies and seniority rewards invites

22. MARGOLIS [1995a] demonstrates these thought experiments in more detail.

further research, both theoretical and empirical, to attempt to explain the sources of this heterogeneity in a convincing manner.

• References

ABOWD, J. M., KRAMARZ, F., MARGOLIS, D. N. (1994). – "High Wage Workers and High Wage Firms", *NBER Working Paper No. 4917*.

ABRAHAM, K. G., FARBER, H. S. (1987). – "Job Duration, Seniority and Earnings", *American Economic Review*, 77.

ABRAHAM, K. G., FARBER, H. S. (1988). – "Returns to Seniority in Union and Nonunion Jobs: A New Look at the Evidence", *Industrial and Labor Relations Review*, 42 (1), pp. 3-19.

ALTONJI, J. G., SHAKOTKO, R. A. (1987). – "Do Wages Rise with Job Seniority?", *Review Of Economic Studies*, 54, pp. 437-459.

BAKER, G., GIBBS, M., HOLMSTROM, B. (1994). – "The Wage Policy of a Firm", *Quarterly Journal of Economics*, 109 (4), pp. 921-955.

BEAUDRY, P., DINARDO, J. (1991). – "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data", *Journal of Political Economy*, 99 (4), pp. 665-688.

BECKER, G. S. (1993). – *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Third ed., Chicago: NBER.

HARRIS, M., HOLMSTROM, B. (1982). – "A Theory of Wage Dynamics", *Review of Economic Studies*, pp. 315-333.

HAUSMAN, J. A. (1978). – "Specification Tests in Econometrics", *Econometrica*, 46 (6), pp. 1251-1271.

HSIAO, C. (1986). – *Analysis of Panel Data*, Cambridge: Cambridge University Press.

JOUANOVIC, B. (1979). – "Job Matching and the Theory of Turnover", *Journal of Political Economy*, 87 (5), pp. 972-990.

MACLEOD, W. B., MALCOMSON, J. M. (1988). – "Reputation and Hierarchy in Dynamic Models of Employment", *Journal of Political Economy*, 96 (4), pp. 832-854.

MARGOLIS, D. N. (1995). – "Cohort Effects and Estimating Heterogenous Returns to Seniority", *mimeo*, Université de Montréal.

MARGOLIS, D. N. (1995). – "Firm Heterogeneity and Workers Self-Selection Bias Estimated Returns to Seniority", Université de Montréal, Département de Sciences Économiques, Working Paper.

MURPHY, K. M., WELCH, F. (1992). – "The Structure of Wages", *Quarterly Journal of Economics*, 107 (428), pp. 285-326.

SALOP, J., SALOP, S. (1976). – "Self Selection and Turnover in the Labor Market", *Quarterly Journal of Economics*, 90.

TOPEL, R. (1991). – "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority", *Journal of Political Economy*, 99 (1), pp. 145-176.