2007s-02

Wages, Productivity and Aging

Benoit Dostie

Série Scientifique
Scientific Series

Montréal
Février 2007

CIRANO
Centre interuniversitaire de recherche
en analyse des organisations

Wages, Productivity
and Aging[*]

Benoit Dostie [†]

Résumé / *Abstract*

Dans le présent article, nous évaluons les écarts de salaires et de productivité en fonction de l'âge à l'aide de données canadiennes liées employeur-employé de l'Enquête sur le milieu de travail et les employés pour la période 1999-2003. Les données obtenues au sujet de l'entreprise ont servi à estimer des fonctions de production en tenant compte du profil d'âge de son effectif. Les données concernant les employés ont été utilisées pour estimer des équations de salaires qui dépendent aussi de l'âge. Les résultats révèlent des profils concaves dans les relations âge-salaire et âge-productivité. Les comparaisons des déterminants des salaires et de la productivité démontrent que la productivité des travailleurs de 55 ans et plus ayant au moins terminé des études de premier cycle est plus faible que leurs salaires. Pour ce qui est des autres groupes, nous constatons que les salaires ne dévient pas de façon statistiquement significative de la productivité.

Mots clés : fixation des salaires, productivité, fonction de production, données liées employeur-employé

In this article, we estimate age based wage and productivity differentials using linked employer-employee Canadian data from the Workplace and Employee Survey 1999-2003. Data on the firm side is used to estimate production functions taking into account the age profile of the firm's workforce. Data on the workers' side is used to estimate wage equations that also depend on age. Results show concave age-wage and age-productivity profiles. Wage-productivity comparisons show that the productivity of workers aged 55 and more with at least an undergraduate degree is lower than their wages. For other groups, we find that wages do not deviate significantly from productivity estimates.

Keywords: *wage determination, productivity, production function, linked employer-employee data*

[†] Institute of applied economics, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7; email: benoit.dostie@hec.ca; fax: 514-340-6469; phone: 514-340-6453; CIRANO, CIRPÉE and IZA.

1 Introduction

Wage differentials based on different levels of schooling or experience are well documented in the labor economics literature. These are generally interpreted as productivity differentials based on an investment model of human capital. Models generally predict that wages increase in the early stages of a career until they reach a plateau, afterwhich, they start do diminish due to human capital depreciation (Mincer (1974)).

There are many reasons, however, to doubt that wage differentials always correspond to differences in productivity. Alternative explanations include incentive-compatible wages (Lazear (1979)), forced saving mechanisms (Frank and Hutchens (1993) and Loewenstein and Sicherman (1991)), efficiency wages (Akerlof and Yellen (1986)), minimization of turnover related costs (Salop and Salop (1976)), specific training (Hutchens (1989)) or wage discrimination.[1]

Obtaining empirical evidence is thus important in order to know which theoretical model best describes the labor market. Moreover, some theoretical models predict that productivity-wage differentials will vary by age. For example, the employer might be using a deferred compensation package (through which workers are paid below productivity at the beginning of their career and above their productivity later) in order to elicit optimal worker effort. If such a model is supported by the data, this would have important consequences for the ability of a country facing an aging population to maintain competitiveness.

In one of the first detailed empirical studies on this topic, Medoff and Abraham (1980) find evidence that the wages of older workers might be higher than their productivity. Corroborating evidence has been found in numerous countries and for different professions using a variety of methodologies, including

---

[1]Discrepancies between wages and productivity could also arise due to labor market imperfections (for example due to minimum wages laws and trade unions, or oligopsonistic wage-setting (Acemoglu and Pischke (1999)), etc.).

Oster and Hamermesh (1998) for economists, Kotlikoff and Gokhale (1992) for workers in the U.S. manufacturing industries, Fair (1994) for American athletes, Crépon, Deniau, and Pérez-Duarte (2003) for France[2] and Haegeland and Klette (1999) for Norway. However, many other studies find the opposite result, for example Mitchell (1990), Smith (1991), Hellerstein and Neumark (1995) and Hellerstein, Neumark, and Troske (1999).

In this article, we estimate age-specific wage and productivity differentials using Canadian data from the Workplace and Employee Survey (WES) 1999-2003. The survey is designed to be representative of all firms operating in Canada and contains detailed information on each firm's production process, organizational practices (and changes in such practices), and human resources policies. Since the survey is linked, there is no need to assign workers to firms using statistical matching methods like Hellerstein, Neumark, and Troske (1999). Also, because the survey is linked, we are able to obtain an external value for a worker's productivity, independent of his or her wage.

We use a methodology that is similar to Hellerstein, Neumark, and Troske (1999) and Aubert and Crépon (2003). More specifically, we use data at the workplace level to estimate production functions taking into account the age composition of the firm's workforce and use data at the employee level to estimate wage equations distinguishing workers based on their age.

However, we improve on their methodologies along several lines. First, we estimate wage equations taking into account both individual and firm unobserved heterogeneity using a mixed model of wage determination (as suggested by Abowd and Kramarz (1999b)). Second, we also control for unobserved time-varying productivity shocks in the production function using a method suggested by Levinsohn and Petrin (2003). Third, we look at the sensitivity of the wage

---

[2]Aubert and Crépon (2003) show that those results are sensitive to the way unobserved productivity shocks are taken into account in the estimation method. Using a different methodology, they cannot reject the hypothesis that wages equal productivity.

productivity-differentials to the measurement of the labor input by the number of employees or hours of work. Finally, we test whether productivity differentials are sensitive to the inclusion of workplace practices and organisational changes in the production function.

We find that wage profiles are not very sensitive to the inclusion of unobserved heterogeneity at the workplace and worker levels. We do find however productivity profiles to be steeper once unobserved productivity shocks are controlled for. Finally, while we find concave profiles for both wage and productivity, our results also show that the productivity of workers aged 55 and more with at least an undergraduate degree is lower than their wages.

The plan of the rest of the paper is as follows. We first describe our methodology in section 2 and present the data and some descriptive statistics in the following section. We describe the results in section 4 and conclude briefly in section 5. All tables are in the appendix.

2   Methodology

Our methodology improves on previous work in two main ways (1) we take into account firm unobserved heterogeneity (in addition to worker unobserved heterogeneity) in the estimation of the wage equation and (2) we also take into account unobserved time-varying productivity shocks using an estimation method suggested by Levinsohn and Petrin (2003) in the estimation of the production function. We describe both models in the following subsections.

2.1 Production function

In order to estimate age-productivity profiles, first consider a Cobb-Douglas production function

$$\log Q_{jt} = \alpha \log L_{jt}^{A} + \beta \log K_{jt} + \gamma F_{jt} + u_{jt} \tag{1}$$

where Q is the value added by firm j at time t, L^A is an aggregate function of different types of workers, K is the capital stock and u the error term. In what follows, types of workers refer to workers of different age, gender and education[3]. F is a matrix of workplace characteristics that are chosen in order to make the specification in (1) as comparable as possible to the specification for the wage equation. F therefore includes industry (13), region (6) and year (4) dummies but also some other characteristics of the workforce like the proportion of workers in each of six occupation categories and three ethnic origins, the proportion of employees that are married, the proportion of immigrants and the proportion of employees covered by a collective bargaining agreement. For each workplace, we observe a representative sample of workers and we use this sample to construct worker proportions mentioned above. We also take advantage of the fact that the WES contains detailed data about workplace practices (6) and organisational changes (14) by including these as controls in the production function.

Let L_{jtk} be the number of workers of type k in firm j at time t, and ϕ_k be their productivity. If we assume that workers of each type are perfectly

---

[3]It would be interesting to also distinguish workers based on other dimensions, for example occupation. However, given the relatively small number of workers that was sampled from each workplace, this would make our constructed proportions too imprecise. We therefore stick to age, gender and education in the analysis that follows.

substitutable, we can write

$$L_{jt}^{A} = \sum_{k=0}^{K} \lambda_k L_{jtk} = \lambda_{j0} L_{jt} + \sum_{k=1}^{K} (\lambda_{jk} - \lambda_{j0}) L_{jtk} \tag{2}$$

where L_{jt} is the total number of workers in the firm λ_0 the productivity of the reference category of workers. If we assume that a worker has the same marginal product accross firm[4], we can rewrite equation (2) as

$$\log L_{jt}^{A} = \log \lambda_0 + \log L_{jt} + \log \left(1 + \sum_{k=1}^{K} \left(\frac{\lambda_k}{\lambda_0} - 1 \right) P_{jkt} \right) \tag{3}$$

where P_{jkt} is the ratio of the number of workers of type k over the total number of employees. We then write the production function as

$$
\begin{aligned}
\log Q_{jt} \;\; = \;\; & \alpha \log \lambda_0 + \alpha \log L_{jt} + \\
& + \alpha \log \left(1 + \sum_{k=1}^{K} \left(\frac{\lambda_k}{\lambda_0} - 1 \right) P_{jkt} \right) + \beta \log K_{jt} + \gamma F_{jt} + u_{jt} \;\; (4)
\end{aligned}
$$

As Hellerstein, Neumark, and Troske (1999), we distinguish three age groups: less then 35, between 35 and 55, and over 55. As to education, we distinguish workers based on whether they have at least an undergraduate degree or not. Therefore, workers are thus separated in 7 categories (men and women (M and W); young, middle age or old (Y, P and O); with or without a degree (D, N)). If we take young male workers without a degree as our reference category, we can write:

---

[4]Hellerstein, Neumark, and Troske (1999) find that their productivity differentials estimates are robust to that assumption.

$$\log L^A_{jt} \quad = \quad \log \lambda_0 + \log L_{jt} +$$

$$+ \log \begin{pmatrix} 1 + \gamma_{MYD}\frac{L_{MYDjt}}{L_{jt}} + \gamma_{MPN}\frac{L_{MPNjt}}{L_{jt}} + \gamma_{MPD}\frac{L_{MPDjt}}{L_{jt}} \\ \gamma_{MON}\frac{L_{MONjt}}{L_{jt}} + \gamma_{MOD}\frac{L_{MODjt}}{L_{jt}} + \\ \gamma_{WYN}\frac{L_{WYNjt}}{L_{jt}} + \gamma_{WYD}\frac{L_{WYDjt}}{L_{jt}} + \\ \gamma_{WPN}\frac{L_{WPNjt}}{L_{jt}} + \gamma_{WPD}\frac{L_{WPDjt}}{L_{jt}} + \\ \gamma_{WON}\frac{L_{WONjt}}{L_{jt}} + \gamma_{WOD}\frac{L_{WODjt}}{L_{jt}} + \end{pmatrix} \quad (5)$$

where γ equal $(\lambda/\lambda_0 - 1)$. Since $\log(1 + x) \simeq x$, we can approximate this by

$$\log L^A_{jt} \quad = \quad \log \lambda_0 + \log L_{jt} + \gamma_{MYD}\frac{L_{MYDjt}}{L_{jt}} +$$

$$\gamma_{MPN}\frac{L_{MPNjt}}{L_{jt}} + \gamma_{MPD}\frac{L_{MPDjt}}{L_{jt}} +$$

$$\gamma_{MON}\frac{L_{MONjt}}{L_{jt}} + \gamma_{MOD}\frac{L_{MODjt}}{L_{jt}} +$$

$$\gamma_{WYN}\frac{L_{WYNjt}}{L_{jt}} + \gamma_{WYD}\frac{L_{WYDjt}}{L_{jt}} +$$

$$\gamma_{WPN}\frac{L_{WPNjt}}{L_{jt}} + \gamma_{WPD}\frac{L_{WPDjt}}{L_{jt}} +$$

$$\gamma_{WON}\frac{L_{WONjt}}{L_{jt}} + \gamma_{WOD}\frac{L_{WODjt}}{L_{jt}} \quad (6)$$

We call this specification the "complete" model. If we impose the following restrictions: $\gamma_{MYD} = \gamma_D$, $\gamma_{MPN} = \gamma_P$, $\gamma_{MPD} = \gamma_P \cdot \gamma_D$, $\gamma_{MON} = \gamma_O$, $\gamma_{MOD} = \gamma_O \cdot \gamma_D$, $\gamma_{WYD} = \gamma_W$, $\gamma_{WYD} = \gamma_W \cdot \gamma_D$, $\gamma_{WPN} = \gamma_W \cdot \gamma_P$, $\gamma_{WPD} = \gamma_W \cdot \gamma_P \cdot \gamma_D$, $\gamma_{WON} = \gamma_W \cdot \gamma_O$, $\gamma_{WOD} = \gamma_W \cdot \gamma_O \cdot \gamma_D$, we can write a more parsimonious specification as

$$\log L^A_{jt} = \log \lambda_0 + \log L_{jt} + \gamma_W \frac{L_{Wjt}}{L_{jt}} + \gamma_P \frac{L_{Pjt}}{L_{jt}} + \gamma_O \frac{L_{Ojt}}{L_{jt}} + \gamma_D \frac{L_{Djt}}{L_{jt}} \quad (7)$$

We call this last specification the "restricted" model. Substituting (2) in (4)

gives us the restricted production function:

$$
\begin{aligned}
\log Q_{jt} \quad \simeq \quad & \beta_0 + \alpha \log L_{jt} + \beta \log K_{jt} + \alpha\gamma_W \frac{L_{Wjt}}{L_{jt}} + \alpha\gamma_P \frac{L_{Pjt}}{L_{jt}} + \\
& \alpha\gamma_O \frac{L_{Ojt}}{L_{jt}} + \alpha\gamma_D \frac{L_{Djt}}{L_{jt}} + \gamma F_{jt} + u_{jt}
\end{aligned}
\tag{8}
$$

and the complete model follows from the substitution of (6) in (4) :

$$
\begin{aligned}
\log Q_{jt} \quad \simeq \quad & \beta_0 + \alpha \log L_{jt} + \beta \log K_{jt} + \alpha\gamma_{MYD} \frac{L_{MYDjt}}{L_{jt}} + \\
& \alpha\gamma_{MPN} \frac{L_{MPNjt}}{L_{jt}} + \alpha\gamma_{MPD} \frac{L_{MPDjt}}{L_{jt}} + \\
& \alpha\gamma_{MON} \frac{L_{MONjt}}{L_{jt}} + \alpha\gamma_{MOD} \frac{L_{MODjt}}{L_{jt}} + \\
& \alpha\gamma_{WYN} \frac{L_{WYNjt}}{L_{jt}} + \alpha\gamma_{WYD} \frac{L_{WYDjt}}{L_{jt}} + \\
& \alpha\gamma_{WPN} \frac{L_{WPNjt}}{L_{jt}} + \alpha\gamma_{WPD} \frac{L_{WPDjt}}{L_{jt}} + \\
& \alpha\gamma_{WON} \frac{L_{WONjt}}{L_{jt}} + \alpha\gamma_{WOD} \frac{L_{WODjt}}{L_{jt}} + \gamma F_{jt} + u_{jt}
\end{aligned}
\tag{9}
$$

where β_0 is a constant term that incorporates $\alpha \log \lambda_0$ and δ is a vector of parameters.

Note that coefficient estimates of equations (8) and (9) will be biased if input choices in the production function are correlated to unobserved productivity shocks (Griliches and Mairesse (1998)). Profit maximizing firms will respond to a positive shock by increasing production, which requires more input. In a similar manner, negative productivity shocks will lead firms to lower their production level. Among the studies using production function to estimate productivity differentials, only Aubert and Crépon (2003) take this problem into account.[5]

Many methods have been proposed to overcome this problem. A popular

---

[5]However, Hellerstein, Neumark, and Troske (1999) do use instrumental variables for materials when estimating a production function on gross output.

estimation strategy relies on system generalized method of moments methods as first proposed by Blundell and Bond (1999). This is the method used by Aubert and Crépon (2003) who find that taking into account unobserved productivity shocks completely reverse the conclusion of Crépon, Deniau, and Pérez-Duarte (2003). However, Gorodnichenko (2006) shows that the Blundell and Bond estimator is in general weakly identified. A comparison of the coefficient estimates between Aubert and Crépon (2003) and Crépon, Deniau, and Pérez-Duarte (2003) shows, as expected, that the productivity differentials are less precisely estimated in the former case.

To correct for endogenous input choices, we thus prefer to use a two-stages estimation method suggested by Levinsohn and Petrin (2003). The idea of their estimator is to invert the demand function for capital and materials to infer a value for the unobserved productivity shock. The estimated productivity shock is then used as a regressor in the production function. See Levinsohn and Petrin (2003) for complete details. We note that the method assumes that the inversion function is non stochastic. If this assumption is violated, estimates will be biased (as argued by Bond and Soderborm (2005), Ackerberg, Caves, and Frazer (2003) and Basu (1999)[6]). However, Gorodnichenko (2006) provides Monte-Carlo evidence showing that the resulting bias is likely to be smaller than the bias from OLS methods, at least in the case of returns to scale estimation. With this caveat in mind, we present two sets of results using OLS and Levinsohn and Petrin (2003) methods.

2.2 Wage equations

Turning to the estimation of the relationship between age and wages, it is possible to use two approaches: wage regressions at the worker level or payroll

---

[6]Alternative estimation strategies proposed by Ackerberg, Caves, and Frazer (2003) and Olley and Pakes (1996) rely on similar assumptions and will share this bias.

regressions. Crépon, Deniau, and Pérez-Duarte (2003) and Hellerstein, Neumark, and Troske (1999) estimate payroll equations for two reasons (1) they enable joint estimation of payroll and production function equations, and therefore yield a direct test of the hypothesis that wages equal productivity for each age group; and (2) they argue that the simultaneous model minimizes the impact of unobserved factors on productivity and wages.

However, an aggregate approach to estimate age-based wage-differentials cannot take into account unobserved heterogeneity at the worker level. This could be important if labor attachment varies by age according to unobserved productivity differences between workers. Therefore, in the analysis that follows, we will favor the disaggregated approach.[7]

In order to take into account both firm and workplace heterogeneity in our model of wage determination, we use a two-factor analysis of covariance with repeated observations along the lines of Abowd and Kramarz (1999b):

$$y_{it} = \mu + \mathbf{x}_{it}\boldsymbol{\beta} + \theta_i + \psi_{j(i,t)} + \epsilon_{it} \tag{10}$$

with

$$\theta_i = \alpha_i + \mathbf{u}_i\boldsymbol{\eta} \tag{11}$$

where y_{it} is the (log) wage rate observed for individual $i = 1, ..., N$, at time $t = 1, ..., T_i$. Person effects are denoted by i, firm effects by j (as a function of i and t), and time effects by t. μ is a constant, $\mathbf{x}_{it}$ is a matrix containing demographic information for employee i at time t[8] as well as information concerning the workplace j to which the worker i is linked. Although β and η can be fixed or random, we assume they are fixed in our estimations. All other

---

[7]Thus, our results obtained from wage equations are not directly comparable to Hellerstein, Neumark, and Troske (1999).

[8]In particular, we include information about age, gender and education in a consistent manner with equations (8) and (9) in order to evaluate wage-productivity differentials.

effects are random. Personal heterogeneity (θ_i) is a measure of unobserved (α_i) and observed $(\mathbf{u}_i \boldsymbol{\eta})$ human capital and follows the worker from firm to firm. Employer heterogeneity (ψ_j) is a measure of firm-specific compensation policies and is paid to all workers of the same firm[9]. ϵ_{it} is the statistical residual.

In full matrix notation, we have

$$y = X\beta + U\eta + D\alpha + F\psi + \epsilon \tag{12}$$

where: y is the $N^* \times 1$ vector of earnings outcomes; X is the $N^* \times q$ matrix of observable time-varying characteristics including the intercept; β is a $q \times 1$ parameter vector; U is the $N^* \times p$ matrix of time invariant person characteristics; η is a $p \times 1$ parameter vector; D is the $N^* \times N$ design matrix of the unobserved component for the person effect; α is the $N \times 1$ vector of person effects; F is the $N^* \times J$ design matrix of the firm effects; ψ is the $J \times 1$ vector of pure firm effects; and ϵ is the $N^* \times 1$ vector of residuals.

Estimation of (12) on large-scale data sets has been achieved by Abowd, Kramarz, and Margolis (1999) while treating firm and person effects as fixed. Here we focus on a mixed-model specification for wage determination because the sampling frame does not follow workers moving from firm to firm. When this is the case, parametric assumptions embedded in the mixed model are necessary to distinguish firm and individual unobserved heterogeneity. Therefore, identification of individual and firm random effects comes from the longitudinal and linked aspects of the data as well as from distributional assumptions. For individual effects, identification comes from the repeated observations on each individual over time. Identification of firm effects comes from repeated observations on workers from the same firm. Our choice for a mixed specification is done without loss of generality since it can be shown that the least squares

---

[9]Firm unobserved heterogeneity in productivity is a common factor in many models of wage dispersion, see Mortensen (2003).

estimates of the fixed effects are a special case of the mixed model estimates (see Abowd and Kramarz (1999b)).

We thus assume α and ψ to be distributed normally :

$$
\begin{bmatrix} \alpha \\ \psi \\ \epsilon \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & \Lambda \end{bmatrix} \right) \tag{13}
$$

where

$$
\Lambda = \begin{bmatrix} \Sigma_1 & 0 & \ldots & & 0 \\ \ldots & \ldots & & & \ldots \\ 0 & \ldots & \Sigma_i & \ldots & 0 \\ \ldots & & & \ldots & \ldots \\ 0 & & \ldots & 0 & \Sigma_N \end{bmatrix}
$$

and

$$
\Sigma_i = V(\epsilon_i)
$$

with

$$
\Omega = \begin{bmatrix} \sigma_\alpha^2 I_N & 0 \\ 0 & \sigma_\psi^2 I_J \end{bmatrix}. \tag{14}
$$

is the matrix of variance components.

Parameters estimates are obtained in two steps. We first use Restricted Maximum Likelihood (REML) methods to get parameter estimates for the variance components in (13). We then solve the mixed equations to get estimates for the other parameters in the full model (12). These steps are discussed in details in Abowd and Kramarz (1999b). However, two important points should be made about the estimates for $\left(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi} \right)$. First, mixed model solutions $\left(\hat{\beta}, \hat{\eta}, \hat{\alpha}, \hat{\psi} \right)$ converge to the least squares solutions as $|\Omega| \to \infty$ (if $\Lambda = \sigma_\epsilon^2 I_{N^*}$). In this sense, the least squares solutions are a special case of the mixed model

solutions. Second, unlike the usual random effects specification considered in the econometric literature, (12) and (13) do not assume that the random effects are orthogonal to the design (X and U) of the fixed effects (β and η), that is we do not assume $X'D = X'F = U'D = U'F = 0$. If this were the case, we could solve for $\hat{\beta}$ and $\hat{\eta}$ independently of $\hat{\alpha}$ and $\hat{\psi}$.

3 Data

We use data from the WES conducted by Statistics Canada annually from the year 1999 to 2003[10]. The survey is both longitudinal and linked in that it documents the characteristics of workers and workplaces over time[11]. The target population for the "*workplace*" component of the survey is defined as the collection of all Canadian establishments who paid employees in March of the year of the survey. The sample comes from the "Business registry" of Statistics Canada, which contains information on every business operating in Canada. The survey, however, does not cover the Yukon, the Northwest Territories and Nunavut. Firms operating in fisheries, agriculture and cattle farming are also excluded.

For the "*employee*" component, the target population is the collection of all employees working, or on paid leave, in the workplace target population. Employees are sampled from an employees list provided by the selected workplaces. For every workplace, a maximum number of 24 employees is selected and for establishments with less than 4 employees, all employees are sampled. In the case of total non-response, respondents are withdrawn entirely from the survey and sampling weights are recalculated in order to preserve representativeness of the sample. The WES selects new employees and workplaces in odd years (at

---

[10]This is a restricted-access data set available in Statistics Canada Research Data Centers (RDC).

[11]Abowd and Kramarz (1999a) classify WES as a survey in which both the sample of workplaces and the sample of workers are cross-sectionally representative of the target population.

every third year for employees and at every fifth year for workplaces). Hence, the survey can only be representative of the whole target population during these re-sampling years.

One limitation of the WES is that the survey does not incorporate a measure of the firm's capital stock. However, Turcotte and Rennison (2003) also use the WES to estimate production function and solve this problem by using industry average capital stock as a proxy for the individual firm's capital stock. We also use this approach in this paper. Industry average capital stocks come from Table 310002 of CANSIM II at Statistics Canada. These capital stocks correspond to net geometric end of year stock for all capital accounts. We then divide these industry averages by the number of firms in each industry to obtain an individual firm's capital stock. We discuss the likely impact of this imputation method below. Table 1 and 2 present descriptive statistics for all variables used in our analysis. Note that it is not possible for confidentiality reasons to show minima and maxima.

4   Results

Table 3 summarizes wage-productivity differentials for the restricted model. The impact of age on wage is shown in the first column while the second and third column show the impact of age on productivity, distinguishing between whether we measure the labor input by the proportion of workers in each age group or the proportion of hours worked by workers in each age group. The first panel shows OLS results while the second panel shows coefficient estimates obtained from the (preferred) mixed model for the wage equation and Levinsohn and Petrin (2003)'s method for the production function.

Our preferred estimates show that both wage and productivity profiles are concave: wage and productivity are both at their highest for the 35-55 age

group and diminish afterward. However, the degree of concavity is sensitive to the way we estimate the model: productivity profiles are steeper if we measure the labor input by the number of employees in each group and wage profiles are also steeper when using OLS. This is probably explained by the fact that more productive workers (due to unobserved reasons) are more likely to stay in the labor force after age 55, but are working fewer hours. It is worth noting that productivity-age profiles do not seem very sensitive to the estimation method we use[12].

Comparing wage and productivity differentials, we cannot reject the hypothesis that wages are equal to productivity. While workers aged between 35 and 55 earn 13% more on average than younger workers, we estimate their productivity to be 15% higher. Similarly, older workers earn on average 12% more but are 13% more productive. Note that we observe a wage-productivity gap when measuring the labor input by the number of workers in each age group. This underlines the importance of correctly measuring labor in the production function.

To assess the robustness of those results, we now turn to coefficient estimates from the complete model where we interact age with gender and education. These results are presented separately for men without (Table 4) or with a degree (Table 5) and women without (Table 6) and without a degree (Table 7). Overall, we should note that standard errors for the age-productivity profiles are much higher than in the restricted model. This is due to the fact that our constructed proportion of workers for each type is less precise due to the small sample of workers selected from each workplace.

Still focusing on our preferred estimates (from the mixed model and LP estimation method with hours) and turning first to men without a degree, we

---

[12]Also, we find those productivity profiles not to be sensitive to the inclusion of explicit controls for workplace practices and organisational changes.

cannot reject, again, the hypothesis that wages are equal to productivity. While the wages of men aged between 35 and 55 are 13% higher than men aged below 35, their productivity is similarly 15% higher. Productivity of the workers aged 55 and more seems to be slightly higher than their wage (15% versus 10%) but the size of the standard error on the productivity differentials (about 5%) does not allow us to conclude that there is a wage-productivity gap.

However, the coefficient estimates for men with a degree tell a completely different story. We estimate that wages of the workers aged 55 and more are 39% higher than wages of the reference category while their productivity is only 8% higher. This is the opposite of what we observe for workers aged 35 and less. This pattern of wage-productivity differentials is thus consistent with an incentive-compatible wages model where employer defer compensation in order to elicit the optimal effort level from workers (see for example Lazear (1979)). It is interesting to note that these differentials appear only in the case of workers with higher level of education where we expect effort levels to be harder to monitor.

A likely reason why we reach a different conclusion than Hellerstein, Neumark, and Troske (1999) is that their wage-productivity comparisons are based on age-wage differentials estimated from a payroll regression. In fact, if we estimate simple payroll regression by OLS (not shown), we get similar differentials to theirs. This underlines the importance of correctly accounting for unobserved worker and workplace heterogeneity in order to get unbiased age-wage profiles. Our results are similar than Aubert and Crépon (2003) for the restricted model but since they do not interact age with gender and education, we cannot compare our results for the complete model where we do observe a differential.

Tables 6 and 7 present wage-productivity differentials for women, depending

on whether they possess at least an undergraduate degree or not. While results for women without a degree are similar to men, in the case of women with a degree, we cannot reject the hypothesis that wages are equal to productivity. Unfortunately, this seems to be mostly due to the fact that productivity differentials are estimated with much less precision. We note again the importance of taking into account unobserved productivity shocks and measuring the labor input by hours of work: results using OLS show a very large wage-productivity gap for women aged 35 and less. However, this gap mostly disappears once we control for such factors.[13]

4.1   Capital Stock

The estimated coefficient on capital stock (not shown) is close to zero in all specifications. This is common in production function estimation using firm-level micro survey data. Aubert and Crépon (2003), for instance, report capital shares of between 0.13–0.14 and Hellerstein, Neumark, and Troske (1999), report an even lower coefficient of 0.05.

That being said, there are at least two explanations for our low capital stock parameter estimates. First, a large proportion of our sample operates in the service sector, in which physical capital does not play the type of role it would in, say, the manufacturing sector. Second – and this is also a common problem – our imputed capital stock measure is likely to be subject to measurement error. To assess the extent of the potential bias introduced by this problem, we reestimate the model omitting the capital stock variable entirely. The age-productivity differentials were virtually identical in the two models. We conclude from this that our productivity differential estimates are robust to our measure of the capital stock.

---

[13]Hellerstein, Neumark, and Troske (1999) also estimated a very large gap for women aged 35 and less and made the hypothesis that this was due to weaker labor attachment. Our results confirm this hypothesis.

5   Conclusion

In this paper, we provide new evidence on the relationship between wages and productivity across the lifecycle. We use linked employer-employee data to estimate wage equations controlling for the age of the worker and estimate production functions that depend on the age structure of each firm's workforce, and compare results from both specifications. Our framework is thus similar to Hellerstein, Neumark, and Troske (1999) and Aubert and Crépon (2003). However, we improve the estimation of wage equations by taking into account both individual and firm unobserved heterogeneity using a mixed model of wage determination (as suggested by Abowd and Kramarz (1999b)). We also control for unobserved time-varying productivity shocks in the production function using a method suggested by Levinsohn and Petrin (2003).

The data used come from the WES 1999-2003 from Statistics Canada. Since the survey is linked, there is no need to assign workers to firms using statistical matching methods like to Hellerstein, Neumark, and Troske (1999). Moreover, the survey is designed to be representative of all firms operating in Canada. We have information on each workplace's production process, organizational practices (and changes in such practices), and human resources policies.

We find evidence that wages of men with at least an undergraduate degree aged 35 and less are lower than their productivity while the reverse is true for men aged 55 and more, a pattern coherent with deferred compensation models. This is a particularly worrying results with respect to the current aging of the workforce. Moreover, the presence of back loading of compensation for this group is likely to translate into fewer hiring opportunities for older workers. In fact, building on the recognition that many workplaces employ older workers but do not hire older workers, many studies (for example Daniel and Heywood (2007) and Hutchens (1988)) find such a link between deferred compensation

and hiring opportunities.

We should note that in all our specifications, age-productivity differentials are estimated with much less precision than age-earning differentials. This is probably due to the fact that the different age groups in the production function are computed using a sample of workers from each firm. One should also note that we distinguish workers only based on age, gender and education. Another important distinction is occupation. For example, it might be important to distinguish workers in managerial positions from workers in production positions. However, our sample of workers from each firm is not big enough to allow such fine distinctions[14].

Finally, all results depend on whether our method for the imputation of the capital stock is realistic or not. Having the right capital stock is important because productivity differentials are computed based on parameters for different age group and on the coefficient on labor (α) in the production function. A bias in this coefficient will translate to a bias in our age-productivity differentials.

References

Abowd, J. M. and F. Kramarz (1999a). The analysis of labor markets using matched employer-employee data. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, vol 3B*, Chapter 40, pp. 2629–2710. Elsevier Science North Holland.

Abowd, J. M. and F. Kramarz (1999b). Econometric analyses of linked employer-employee data. *Labour Economics 6*(1), 53–74.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

---

[14]However, comparisons of point estimates obtained when we do allow such a distinction suggest that the pattern of wage-productivity differentials found in this paper holds accross occupations

Acemoglu, D. and S. Pischke (1999). Beyond Becker: Training in imperfect labor markets. *The Economic Journal 119*, F112–F142.

Ackerberg, D., K. Caves, and G. Frazer (2003). Structural identification of production functions. Working paper, University of Arizono, UCLA and University of Toronto.

Akerlof, G. A. and J. L. Yellen (1986). *Efficiency Wage Models of the Labor Market*. Cambridge University Press.

Aubert, P. and B. Crépon (2003). Âge, salaire et productivité : La productivité des salariés décline-t-elle en fin de carrière? *Économie et Statistique 368*, 95–119.

Basu, S. (1999). Discussion on "Estimating production function using intermediate inputs to control for unobservables" by A. Petrin and J. Levinsohn. Working paper, NBER Productivity Program Meeting.

Blundell, R. and S. Bond (1999). GMM estimation with persistent panel data: An application to production functions. Working Paper 99/04, Institute for Fiscal Studies.

Bond, S. and M. Soderborm (2005). Adjustment costs and the identification of Cobb-Douglas production functions. Working Paper 05/04, Institute for Fiscal Studies.

Crépon, B., N. Deniau, and S. Pérez-Duarte (2003). Productivité et salaire des travailleurs âgés. *Revue française d"économie 18*(1), 157–185.

Daniel, K. and J. S. Heywood (2007). The determinants of hiring older workers: UK evidence. *Labour Economics 14*, 35–51.

Fair, R. (1994). How fast do old men slow down. *Review of Economics and Statistics 76*(1), 103–118.

Frank, R. H. and R. M. Hutchens (1993). Wages, seniority, and the demand

for rising consumption profiles. *Journal of Economics Behavior and Organization 21*, 251–76.

Gorodnichenko, Y. (2006). Using firm optimization to evaluate and estimate returns to scale. Working paper, University of Michigan.

Griliches, Z. and J. Mairesse (1998). Production functions: The search for identification. In S. Strom (Ed.), *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, pp. 169–203. Cambridge University Press.

Haegeland, T. and T. Klette (1999). Do higher wages reflect higher productivity? Education, gender and experience premiums in a matched plant-worker data set. In J. Haltiwanger, J. Lane, J. Spletzer, J. Theeuwes, and K. Troske (Eds.), *The Creation and Analysis of Employer-Employee Matched Data*, pp. 231–259. Elsevier Science, North-Holland.

Hellerstein, J. K. and D. Neumark (1995). Are earnings profiles steeper than productivity profiles? Evidence from Israeli firm data. *Journal of Human Resources 30*(1), 89–112.

Hellerstein, J. K., D. Neumark, and K. Troske (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equation. *Journal of Labor Economics 17*(3), 409–446.

Hutchens, R. M. (1988). Do job opportunities decline with age? *Industrial and Labor Relations Review 42*, 89–99.

Hutchens, R. M. (1989). Seniority, wages and productivity: A turbulent decade. *Journal of Economic Perspectives 2*(4), 49–64.

Kotlikoff, L. J. and J. Gokhale (1992). Estimating a firm's age-productivity using the present value of workers' earnings. *Quarterly Journal of Economics 107*(4), 1215–1242.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy 87*, 1261–1284.

Levinsohn, J. and A. Petrin (2003). Estimating production function using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–342.

Loewenstein, G. and N. Sicherman (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics 9*, 67–84.

Medoff, J. L. and K. G. Abraham (1980). Experience, performance, and earnings. *Quarterly Journal of Economics 95*(4), 703–736.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: NBER Press.

Mitchell, O. (1990). Aging, job satisfaction, and job performance. In I. Bluestone, R. Montgomery, and J. Owen (Eds.), *An Aging Workforce*, pp. 242–272. Wayne State University Press, Detroit, MI.

Mortensen, D. T. (2003). *Wage Dispersion : Why Are Similar Workers Paid Differently? (Zeuthen Lectures)*. The MIT Press.

Olley, G. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*, 1263–1297.

Oster, S. and D. Hamermesh (1998). Aging and productivity among economists: Note. *Review of Economics and Statistics 80*(1), 154–56.

Salop, J. and S. C. Salop (1976). Self-selection and turnover in the labor market. *Quarterly Journal of Economics 90*(4), 619–627.

Smith, S. (1991). Ending mandatory retirement in the arts and sciences. *American Economic Review 81*(2), 106–110.

Turcotte, J. and L. W. Rennison (2003). Productivity and wages: Measuring the effect of human capital and technology use from linked employer-

employee data. Working paper, Department of Finance. Economic and Fiscal Policy Branch.

A   Tables

Table 1: Descriptive statistics - Employees

	1999	
	Mean	Std Dev
ln(Wage)	2.778	0.521
Highest completed degree		
Less then high school	0.107	0.309
High school	0.175	0.380
Industry training	0.053	0.162
Trade or vocational diploma	0.088	0.283
Some college	0.104	0.305
Completed college	0.181	0.385
Some university	0.077	0.266
Teacher's college	0.002	0.049
University certificate	0.018	0.132
Bachelor degree	0.130	0.337
University certificate ($>$ bachelor)	0.019	0.135
Master's degree	0.031	0.174
Degree in medicine, dentistry, etc.	0.008	0.092
Earned doctorate	0.006	0.078
Experience	16.167	10.714
Seniority	8.517	8.206
Black	0.011	0.104
Other races	0.280	0.449
Women	0.521	0.500
Married	0.566	0.496
Immigrant	0.175	0.380
Years since immigration	3.988	10.181
Union	0.279	0.449
Ptime	0.051	0.220
Occupations		
Manager	0.151	0.358
Professional	0.162	0.368
Technician	0.390	0.488
Marketing/sales	0.084	0.277
Clerical/administrative	0.140	0.347
Production w/o certificate	0.074	0.262
Number of employees:	23540	

Table 2: Summary statistics - Workplaces

	1999	
	Mean	Std Dev.
Value added ($)	1 235 394	2.12E-07
Number of employees	12.825	54.418
Capital stock ($)	46749	71770
Gross payroll ($)	406127	2759789
Union	0.046	0.181
Labor force		
Proportion of men	0.400	0.404
Proportion aged between 35 and 55	0.478	0.388
Proportion aged 55 and over	0.211	0.362
Proportion with a degree	0.244	0.375
Number of workplaces:	5499	

Table 3: Wage-productivity differentials - Restricted model - 99-03

	OLS					
	Wage		Prod.		Prod.(Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
$[35 \leq \text{Age} < 55]$	0.147***	1.16	0.124**	1.13	0.158***	1.16
	(0.006)		(0.051)		(0.051)	
$[55 \leq \text{Age}]$	0.092***	1.10	-0.003	1.00	0.167**	1.17
	(0.009)		(0.060)		(0.071)	
	Mixed and LP					
	Wage		Prod.		Prod.(Hours)	
$[35 \leq \text{Age} < 55]$	0.125***	1.13	0.131***	1.15	0.136***	1.15
	(0.003)		(0.029)		(0.024)	
$[55 \leq \text{Age}]$	0.110***	1.12	-0.089***	0.90	0.113**	1.13
	(0.004)		(0.033)		(0.049)	

Reference category: $[\text{Age} < 35]$
Statistically significant at: *=10%; **=5%; ***=1%
Both equations include controls for industry (14), year (5) and region (7).
Wage equation includes controls for union status, part time, black, other race, married, immigrant, years since migration, occupations (6) and a constant. N=99425.
Production function includes controls for proportion of workers in each of six occupation category and three ethnic origins, the proportion of employees that are married, of immigrants, and of employees covered by an union, controls for organisation change and workplace practices. It also includes a constant, the log number of employees and the log of the imputed capital stock. N=20593 (OLS), N=20225 (LP).

Table 4: Differentials - [Men] and [No degree]

	OLS					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[35 ≤ Age < 55]	0.175***	1.19	0.043	1.04	0.011	1.01
	(0.010)		(0.066)		(0.069)	
[55 ≤ Age]	0.095***	1.10	0.135	1.14	0.271***	1.30
	(0.015)		(0.105)		(0.070)	
	Mixed and LP					
	Wage		Prod.		Prod. (Hours)	
[35 ≤ Age < 55]	0.122***	1.13	0.182***	1.20	0.135***	1.15
	(0.004)		(0.042)		(0.037)	
[55 ≤ Age]	0.098***	1.10	0.194***	1.22	0.135**	1.15
	(0.006)		(0.067)		(0.063)	

Reference category: [Age < 35] and [Men] and [No degree]
Statistically significant at: *=10%; **=5%; ***=1%
Both equations include controls for industry (14), year (5) and region (7).
Wage equation includes controls for union status, part time, black, other
race, married, immigrant, years since migration, occupations (6) and a
constant. N=99425.
Production function includes controls for proportion of workers in each
of six occupation category and three ethnic origins, the proportion
of employees that are married, of immigrants, and of employees covered
by an union, controls for organisation change and workplace practices.
It also includes a constant, the log number of employees and the log of
the imputed capital stock. N=20593 (OLS), N=20225 (LP).

Table 5: Differentials - [Men] and [Degree]

	OLS					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	0.207***	1.23	0.540***	1.56	0.463***	1.48
	(0.015)		(0.124)		(0.131)	
[35 ≤ Age < 55]	0.394***	1.48	0.449***	1.47	0.388***	1.40
	(0.015)		(0.099)		(0.100)	
[55 ≤ Age]	0.397***	1.49	0.031	1.03	-0.029	0.97
	(0.028)		(0.199)		(0.208)	
	Mixed and LP					
	Wage		Prod.		Prod. (Hours)	
[Age < 35]	0.108***	1.11	0.469***	1.52	0.391***	1.43
	(0.007)		(0.081)		(0.068)	
[35 ≤ Age < 55]	0.304***	1.36	0.535***	1.60	0.452***	1.50
	(0.006)		(0.071)		(0.056)	
[55 ≤ Age]	0.331***	1.39	0.166	1.05	0.072	1.08
	(0.011)		(0.146)		(0.151)	

Reference category: [Age < 35] and [Men] and [No degree]
Statistically significant at: *=10%; **=5%; ***=1%
Both equations include controls for industry (14), year (5) and region (7).
Wage equation includes controls for union status, part time, black, other
race, married, immigrant, years since migration, occupations (6) and a
constant. N=99425.
Production function includes controls for proportion of workers in each
of six occupation category and three ethnic origins, the proportion
of employees that are married, of immigrants, and of employees covered
by an union, controls for organisation change and workplace practices.
It also includes a constant, the log number of employees and the log of
the imputed capital stock. N=20593 (OLS), N=20225 (LP).

Table 6: Differentials - [Women] and [No degree]

	OLS					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	-0.105***	0.90	-0.340***	0.65	-0.373***	0.61
	(0.011)		(0.083)		(0.088)	
[35 ≤ Age < 55]	-0.010	0.99	-0.063	0.93	-0.081	0.92
	(0.010)		(0.065)		(0.070)	
[55 ≤ Age]	-0.052***	0.95	-0.001	1.00	-0.016	0.98
	(0.014)		(0.107)		(0.110)	
	Mixed and LP					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	-0.117***	0.89	-0.113**	0.87	-0.164***	0.82
	(0.005)		(0.050)		(0.037)	
[35 ≤ Age < 55]	-0.027***	0.97	0.058	1.06	0.010	1.01
	(0.005)		(0.041)		(0.036)	
[55 ≤ Age]	-0.041***	0.96	0.082	1.09	0.033	1.04
	(0.007)		(0.078)		(0.070)	

Reference category: [Age < 35] and [Men] and [No degree]
Statistically significant at: *=10%; **=5%; ***=1%
Both equations include controls for industry (14), year (5) and region (7).
Wage equation includes controls for union status, part time, black, other
race, married, immigrant, years since migration, occupations (6) and a
constant. N=99425.
Production function includes controls for proportion of workers in each
of six occupation category and three ethnic origins, the proportion
of employees that are married, of immigrants, and of employees covered
by an union, controls for organisation change and workplace practices.
It also includes a constant, the log number of employees and the log of
the imputed capital stock. N=20593 (OLS), N=20225 (LP).

Table 7: Differentials - [Women] and [Degree]

	OLS					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	0.066***	1.07	0.008	1.08	-0.084	0.91
	(0.015)		(0.133)		(0.141)	
[35 ≤ Age < 55]	0.266***	1.30	0.199	1.21	0.166	1.17
	(0.013)		(0.128)		(0.116)	
[55 ≤ Age]	0.190***	1.21	0.259	1.27	0.178	1.18
	(0.029)		(0.177)		(0.168)	
	Mixed and LP					
	Wage		Prod.		Prod. (Hours)	
	Coef.	Ratio	Coef.	Ratio	Coef.	Ratio
[Age < 35]	0.003***	1.00	0.206***	1.23	0.161*	1.18
	(0.007)		(0.073)		(0.090)	
[35 ≤ Age < 55]	0.162***	1.18	0.354***	1.39	0.271***	1.30
	(0.007)		(0.073)		(0.070)	
[55 ≤ Age]	0.139***	1.15	0.269	1.30	0.218	1.24
	(0.014)		(0.176)		(0.193)	

Reference category: [Age < 35] and [Men] and [No degree]
Statistically significant at: *=10%; **=5%; ***=1%
Both equations include controls for industry (14), year (5) and region (7).
Wage equation includes controls for union status, part time, black, other
race, married, immigrant, years since migration, occupations (6) and a
constant. N=99425.
Production function includes controls for proportion of workers in each
of six occupation category and three ethnic origins, the proportion
of employees that are married, of immigrants, and of employees covered
by an union, controls for organisation change and workplace practices.
It also includes a constant, the log number of employees and the log of
the imputed capital stock. N=20593 (OLS), N=20225 (LP).