ELSEVIER

Labour Economics 9 (2002) 375–404

LABOUR ECONOMICS

www.elsevier.com/locate/econbase

Matching, human capital, and the covariance structure of earnings

Daniel Parent *,[1]

Department of Economics, McGill University, 855 Sherbrooke St. W., Montreal, Quebec, Canada H3A 2T7

Received 1 January 2001; received in revised form 1 September 2001; accepted 1 November 2001

Abstract

This paper tests the theory of job matching and the theory of human capital by examining the covariance structure of residuals from a typical Mincer log earnings equation using methods of moments techniques. Job matching theory predicts that we should observe an eventual decrease in the contribution of the job-match component in the residual variance as workers acquire tenure on the job. This prediction is mildly supported by the data. On the other hand, human capital theory predicts a trade-off between job-specific intercept and slope parameters. This prediction, which is not shared by the theory of matching, is strongly supported by the data. This is especially true for men with at least a high school degree.

Keywords: Matching; Firm-specific human capital; Generalized method of moments

1. Introduction

Two competing explanations for the existence of a positive return to tenure are human capital and matching. Both theories predict that the conditional mean of wages should rise with tenure.[2] In this paper, I attempt to distinguish between these two theories by

* Department of Economics, McGill University, 855 Sherbrooke St. W., Montreal, Quebec, Canada H3A 2T7.

E-mail address: daniel.parent@mcgill.ca (D. Parent).

[1] CIRANO, Montreal, Canada.

[2] See Becker (1975)'s development and exposition of the theory of human capital and Jovanovic (1979a,b, 1984) for the job matching argument. Mortensen (1988) incorporates both human capital and matching and derives the optimal separation strategies. For an analysis of worker mobility patterns in the United States and their relationship with both models, see also Farber (1999).

PII: S0927-5371(02)00008-8

focusing on their implications for the covariance structure of earnings. Using data from the National Longitudinal Survey of Youth (NLSY) over the period 1979–1996, I find strong evidence of a negative correlation between job-specific wage growth and the job-specific intercept, a prediction of the theory of human capital that is not shared by the theory of job matching.

The existence of a return to employer tenure has been a subject of controversy over the last 15 years. Although ordinary least-squares estimates indicate a positive relationship between wages and tenure (e.g. Mincer and Jovanovic, 1981), endogeneity problems have led researchers to develop estimation methods that account for the fact that firm seniority is likely to be correlated with unobservable factors, such as the quality of the worker/firm match. Using data from the Panel Study of Income Dynamics (PSID), Abraham and Farber (1987), Altonji and Shakotko (1987), and Altonji and Williams (1997) find no evidence of a significant return to tenure, while Topel (1991) finds large returns. With data from the National Longitudinal Survey of Young Men, Marshall and Zarkin (1987) also find quite a small tenure effect once they empirically control for the selection process by which only acceptable wage offers on the current job are ever observed. Other authors have focused on the importance of industry-specific capital as a factor of wage growth. With data from the Displaced Worker Surveys, Neal (1995) finds that tenure with the predisplacement employer is positively correlated with the wage earned in the post-displacement job only for those workers who stay in the same industry, a result that is difficult to reconcile with the view that tenure measures only the accumulation of firm-specific skills. Using data from the PSID and the NLSY, Parent (2000) finds that by creating a measure of industry tenure and adding this measure to the log-wage equation, the firm tenure effect all but disappears.

At a theoretical level, MacLeod and Malcomson (1993) find, in a simple model with competitive contract formation, that a positive or negative return to seniority can occur, depending on the nature of specific investments and the structure of market returns for human capital.

Thus, we can see that finding no return to firm seniority when estimating a wage regression cannot necessarily be interpreted as evidence against the importance of firm-specific investments in an employment relationship. For one thing, it might simply reflect the fact that firms are paying the full costs of such investments as well as reaping the full returns. Hence, working with the first conditional moments of wages may not be the appropriate or more convincing way of assessing the relative importance of human capital vs. matching. In previous work, Hause (1980) modeled the second moments of wages, their variances and covariances, to test the prediction of the general human capital model that those who invest more should have both a steeper slope and a lower intercept in their wage profile. This prediction implies that the covariance between these two individual-specific parameters should be negative. In this paper, I extend Hause's approach by introducing jobs into the analysis. It is then possible to isolate a similar key prediction of human capital theory that there should be a trade-off between the job-specific intercept and the tenure slope. Other things being equal, those who start out with a lower salary invest more in human capital and consequently should have a steeper slope. In contrast, the pure theory of matching does not predict such a trade-off within jobs. In fact, in his work on the complementarity between the quality of a match and firm-specific investment, Jovanovic (1979a) shows that a better match should, ceteris paribus, involve more investment.

I find strong evidence of such a trade-off. The only way that the theory of matching can account for the trade-off between the job-specific slope and the intercept is through human capital considerations. All this is not to say that matching is not an important phenomenon. For example, I show that the statistical fit provided by the simplest form of covariance structure implied by the theory of matching represents a significant improvement over a simpler structure that ignores matching altogether. Still, the patterns identified in the data provide strong support to the idea that human capital theory plays an important role in the wage formation process.

A caveat is that this approach to discriminating between matching and human capital models rests on an equalizing differences view of human capital accumulation. Supposing that people differ only in their decision to invest, Rosen (1977) notes that there must be some unobservable factor that underlies this different behavior. This leads to a self-selection problem. Rosen points out that the proportion of variance explained by measurable factors in standard earnings function estimates is but a fraction of the total variance, even for individuals who have apparently similar attributes. By using panel data and by focusing on the restrictions that matching and human capital put on the time pattern of the variance of the residuals from a log-wage regression, one can provide additional control for unobservable individual characteristics while at the same time identify patterns consistent with either one of these theories.[3]

The paper is organized as follows. Section 2 match sketches the main results from the theory of matching with an emphasis on the predictions pertaining to the variance of the job-match component of the residual. Section 3 follows with a discussion of the theory of human capital and the covariance structure of earnings. Section 4 focuses on the empirical implementation of the models discussed in Sections 2 and 3, including discussions of the data and the econometric methodology. Section 5 presents the results, and at the end of the section, I provide a discussion relating the results contained in this paper to the literature on the impact of training on the starting wage of trainees. Section 6 concludes the paper.

2. The theory of matching and its empirical implications

Originating with Stigler's (1961) model of search, economists have developed tools for analyzing the way in which individual units, such as firms or workers, gather and process information. For example, McCall (1970) pioneered the development of the theory of sequential job search in which individuals draw (at a cost) from a wage distribution F and then decide to either accept the offer in hand or to reject it and sample once again. By applying the principle of optimality, each individual will have a decision rule that is characterized by a reservation wage w_R such that any wage offer above this threshold is accepted and all others rejected.

[3] Abowd et al. (1999) find that once unobservable individual characteristics are explicitly taken into account, over 90% of the total variance is explained by the model.

As emphasized by Rothschild (1973), one problem with this model is the existence of the wage distribution itself in equilibrium. If we assume the population of workers to be homogeneous, then no wage less than w_R would be observed and profit-maximizing firms would see no point in offering any wage above w_R. Therefore, the wage distribution would collapse to a trivial one with all the probability mass concentrated at w_R.

The theory of matching offers an answer to this problem by allowing, in an equilibrium context, the existence of both a nontrivial wage distribution and an optimal search strategy by individuals (e.g. Jovanovic, 1979a,b, 1984). The model described below is a discrete-time version of Jovanovic's (1984) model. I will focus mainly on the information-processing part of the model as the predictions we are interested in stem from this learning mechanism.

2.1. The theory in a nutshell

The quality of a match between a worker and a firm can be characterized as being drawn from some known normal distribution. The worker and the firm come from homogeneous populations. The match parameter represents the marginal productivity of the worker in the firm. The theme recurring in this analysis is that both the firm and the worker will gain information on the "true" quality of the match first by merely contacting a potential employer, then by observing the evolution of the worker's output over time. Both sources of information are noise-ridden, and each party is assumed to use Bayes' law to update its beliefs on the true quality of the match.

In the following setup, which borrows heavily from Mortensen (1988), workers already employed receive offers from potential employers in addition to wage offers from their current employer. The latter arise when new information (i.e. an output signal) arrives as to the quality of the match. I also assume that workers who have not yet transited into the labour market receive outside offers at the same frequency as those already employed. While this may not be totally realistic, it leaves unchanged the empirical implications I will focus on. The main assumptions can be summarized as follows:

(1) The prior distribution of θ is $N(\mu,\ 1/\tau_\theta)$, where τ_θ represents the precision of the distribution. This distribution is stable over time. All workers and firms share this prior distribution.

(2) The outcome of the initial screen is denoted as $m = \theta + \varepsilon_m$ where ε_m is distributed $N(0, 1/\tau_m)$. As $\tau_m \to \infty$, all information about the quality of a potential match can be learned merely by contacting an employer. In other words, jobs are *pure search* or *inspection* goods. On the other hand, as $\tau_m \to 0$, nothing can be learned from the match without gaining some experience in it. In that case, jobs are *pure experience* goods.

(3) The output at each period is $x_t = \theta + \varepsilon_t$, with $\varepsilon_t \sim N(0, 1)$ and $E[\varepsilon_t\, \varepsilon_s] = 0$, for all $t \neq s$. At each period t, both parties observe the worker's output and both are equally well informed about the quality of the match. They are assumed to use Bayes' law to update their beliefs on the perceived quality of the match.

(4) All draws from the prior distribution are independent from one another. The quality of the present match provides no information on the quality of potential matches. There is no recall (or if recalls are allowed, the quality of the new match is independent of the old one).

(5) Firms are assumed to pay workers their expected marginal product. This is the equilibrium contract of Jovanovic (1979b) derived under the assumption that firms bid for workers by offering them lifetime contracts, which they are assumed to honor ex post.[4]

(6) $b > 0$ represents the value per period of time spent not employed. It is the income-equivalent of the disutility of work plus any unemployment benefit (or welfare payment).

(7) Wage offers on a specific job and outside offers from potential employers arrive at Poisson rates equal to η and λ, respectively.

(8) Workers maximize the expected present value of their lifetime income.

(9) There is no accumulation of human capital.

Note that the nature of the information acquisition process can be summarized by the parameter η: if $\eta = 0$, the wage on a job is constant throughout the employment relationship, which means that the job is a pure search good. Also, given that the individual always has the choice of not working and enjoy an income stream of b in any given period, workers will not necessarily accept the first job that they are offered.

The time sequence of events is thus the following: (i) a firm and a worker make contact; (ii) upon contact, they both observe a noisy signal m, which provides information on the true value of the match; (iii) the firm makes an offer to the worker on the basis of that information, and the worker decides whether to accept it or not; if he rejects it, he waits for a subsequent offer and the process re-starts at (i); (iv) if he accepts the job offer, the worker collects the wage and both parties update their belief on the quality of the match when they receive new information provided by the first output signal x_1; (v) Following the updating process, the worker, who may or may not have an outside offer w_a in hand, is offered a new wage in the current job and has to decide whether to accept it or to reject it; if he accepts it, the process goes back to stage (iv); if he declines the offer, the worker can either choose nonemployment and collect b and wait for another job offer or, if he has an outside offer w_a in hand, the worker accepts it and switches jobs.

Under essentially the same assumptions, Mortensen (1988) shows that the worker's optimal separation strategy follows a reservation wage policy: he defines the *reservation wage* as the wage offer by the current employer that makes the worker indifferent between continuing the current relationship and transiting to nonemployment (and collect b). This reservation wage is shown to be increasing with tenure. Intuitively, the upside potential of a job decreases with tenure as both the worker and the firm make an increasingly precise assessment of the quality of the match.

Fig. 1 illustrate the case of a worker who stays with her current employer until she reaches period two, at which point she decides to transit to nonemployment. As we can

[4] It can be argued that firms do not have to pay the workers their expected marginal productivity if it deviates from their market-wide productivity. The best option for workers would still be to stay with the same employer provided they are paid their market value. In other words, paying workers at the value of their expected marginal product is not a subgame perfect strategy for the firms. I abstract from this and other contractual considerations that would render the analysis intractable. My strategy is to adopt the simplest version of the theory of matching and to investigate whether the data seem to support it. In any event, a significant portion of the variance of wage residuals is captured by an idiosyncratic job-match component, as we will see below. If wages only reflected the workers' market value, then one would not expect a job-match component to explain part of the variance.

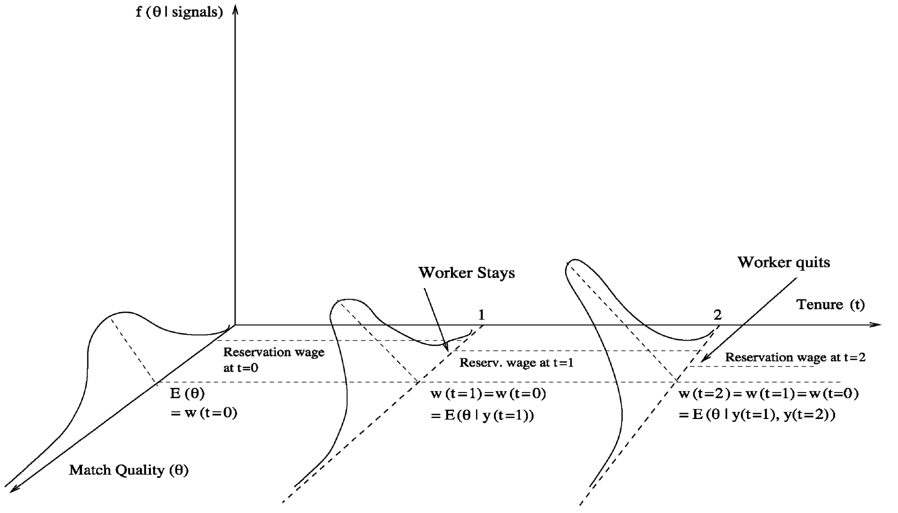

Fig. 1. Worker's choice between employment and nonemployment.

see, even though each new signal essentially "confirms" the a priori information about the expected value of θ and thus leaves the wage unchanged, the variance of the distribution decreases with each new piece of information. This has the effect of reducing the option value of the job and eventually makes the worker quit at $t=2$.

Mortensen also defines the *reservation offer* as the outside wage offer such that the individual is indifferent between his current job and switching to a new employer. For the same reason that the reservation wage increases with tenure, the reservation offer will *decrease* as the worker accumulates seniority, conditional on the wage paid. If, for a given wage, the uncertainty about the potential of a match dissipates, the option value of the job will decrease. This makes the worker more willing to accept a lower wage from a potential employer because the option value of any new job is the same (assuming that jobs are experience goods and that the environment is stationary), irrespective of tenure with the current employer.[5]

Under the stated assumptions, a worker's decision rule can thus be summarized in the following way:

Unemployed worker

• Accept w^* if and only if $w^* \geq w_0^R$ where w^* is the wage offered by an employer and w_0^R is the worker's reservation wage at tenure level zero. Otherwise, reject w^*, collect b and wait for the next offer.

[5] One of Mortensen's main points is that the optimal separation strategies of the matching model and the on-the-job training model are qualitatively the same: the reservation offer decreases with tenure. In the OJT model, this is driven by the assumption that wage increments in a job are decreasing with tenure, a consequence of an optimal investment program in which training intensity is at its highest early on. Workers would become increasingly likely to accept a lower paying job in exchange for the prospect of higher earnings growth at the start of that new job. This point is explored by Topel and Ward (1992).

Employed worker
Suppose the worker has a wage offer of w_t^* from his current employer at time t. Then

- If he does not have an outside offer w_a: transit to nonemployment iff:

  $$w_t^* < w_t^{\mathrm{R}}$$

  where w_t^{R} is the worker's reservation wage at tenure level t.
- If he has an outside offer w_a:

  (i) Transit to nonemployment iff:

  $$w_t^{\mathrm{R}} > \max(w_a, w_t^*)$$

  (ii) Change jobs iff:

  $$w_a \geq \max(w_t^*, w_t^{\mathrm{R}})$$

  (iii) Stay with the same employer iff:

  $$w_t^* \geq \max(w_a, w_t^{\mathrm{R}})$$

The same sort of decision rule is taken into account by Marshall and Zarkin (1987) in estimating the "true" return to tenure: when only acceptable wage offers from the current employer are observed, a simple OLS regression of log wages on tenure will produce upward estimates of the effect of tenure on wages.

Next, I want to characterize the wage structure implied by the theory of matching. The first proposition describes the wage process resulting from assumptions 1–9 and then I focus on how the selection process implied by the fact that good matches last creates a gradual truncation of the wage distribution for stayers.

Proposition 1. *For a worker with no job offers after his initial offer, the wage paid at period t, conditional on the outcome of the screen and on the sequence of outputs is equal to*

$$E\left[\theta \mid m, x_1, x_2 \ldots, x_t\right] = \frac{t\bar{x} + \tau_m m + \tau_\theta \mu}{t + \tau_m + \tau_\theta} \triangleq w_t \tag{1}$$

where $\bar{x} = \frac{1}{t}\sum_{k=1}^{t} x_k$, *with the wage's conditional variance given by*

$$\mathrm{Var}\left[\theta \mid m, x_1, x_2, \ldots, x_t\right] = \frac{1}{t + \tau_m + \tau_\theta} \triangleq S_t \tag{2}$$

The current wage and wage variance conditional on the previous wage is

$$E[w_t | w_{t-1} = w] = w \tag{3}$$

$$\mathrm{Var}[w_t | w_{t-1} = w] = S_t S_{t-1} \tag{4}$$

The initial wage offer distribution is characterized by

$$w_0 \sim N\left(\mu, \frac{\tau_m}{\tau_\theta} S_0\right) \tag{5}$$

In general, the unconditional distribution is characterized by

$$w_t \sim N\left(\mu, \frac{1}{\tau_\theta}\left[1 - S_t^2(t+\tau_\theta)(\tau_m+\tau_\theta)\right]\right) \tag{6}$$

Details are contained in Appendix A. As described in Proposition 1, the variance of the initial wage offer distribution depends in general on the amount of information contained in the initial screening: if jobs are pure experience goods, then every worker is offered the same wage initially; otherwise, the distribution is nondegenerate.

In characterizing the conditional moments of the wage distribution, Proposition 2 makes use of the result in Mortensen (1988), which states that the reservation wage increases with tenure. Consequently, I focus on how the conditional variance changes as the lower threshold represented by the reservation wage keeps increasing. I first do so for the extreme case where jobs are pure search goods so as to highlight the sharp prediction concerning the evolution of the conditional variance of wages for stayers.

Proposition 2. *Let w_t^* be the wage paid on the current job and w_t^R the reservation wage. When jobs are pure search goods, i.e. as $\tau_m \to \infty$, the moments of the truncated distribution are then given by*

$$E\left[w_t^* | w_t^* > w_t^{\mathrm{R}}\right] = \mu + \frac{1}{\sqrt{\tau_\theta}} \frac{\phi\left[\sqrt{\tau_\theta}(w_t^{\mathrm{R}} - \gamma)\right]}{1 - \Phi\left[\sqrt{\tau_\theta}(w_t^{\mathrm{R}} - \gamma)\right]} = \mu + \frac{1}{\sqrt{\tau_\theta}}\lambda(\omega_t) \tag{7}$$

$$\mathrm{Var}\left[w_t^* | w_t^* > w_t^{\mathrm{R}}\right] = \frac{1}{\tau_\theta}\left[1 + \omega_t\lambda(\omega_t) - (\lambda(\omega_t))^2\right] = \frac{1}{\tau_\theta}\left[1 - \delta(\omega_t)\right] \tag{8}$$

where ϕ and Φ are the standard normal density and distribution functions, $\omega_t = \sqrt{\tau_\theta}(w_t^{\mathrm{R}} - \gamma)$,$\lambda(\omega t)$ is the inverse Mills ratio of ω_t, and $\delta(\omega_t) = \lambda(\omega_t)[\lambda(\omega_t) - \omega_t] = d\lambda(\omega_t)/d\omega_t$.

A proof concerning the moments of the truncated normal distribution in Proposition 2 can be found in Johnson and Kotz (1970).

Note that it may be shown that $\omega_t < \lambda(\omega_t)$ and $\delta(\omega_t) \in (0,\ 1)$, $\forall \omega_t \in \mathbf{R}$. Taking a few limits results gives $\lambda(\omega_t) \to 0$ as $\omega_{t-1} \to -\infty$, and $\omega_t - \lambda(\omega_t) \to 0$, as $\omega_t \to +\infty$. Taking derivatives at the limit implies that $\delta(\omega_t) \to 0$ as $\omega_t \to -\infty$, and $\delta(\omega_t) \to 1$ as $\omega_t \to +\infty$, as $\delta(\omega_t)$ is uniformly continuous. The effect of truncation on the variance of the wage distribution is summarized in the following proposition:

Proposition 3. *When jobs are pure search goods, it follows that*

$$\mathrm{Var}\left[w_t^* | w_t^* > w_t^{\mathrm{R}}\right] \to 1 \text{ as } \omega_t \to -\infty \tag{9}$$

$$\mathrm{Var}\left[w_t^* | w_t^* > w_t^{\mathrm{R}}\right] \to 0 \text{ as } \omega_t \to +\infty \tag{10}$$

Furthermore, it may be shown that,

$$\frac{\mathrm{dVar}\left[w_t^*|w_t^* > w_t^{\mathrm{R}}\right]}{\mathrm{d}w_t^{\mathrm{R}}} = \frac{1}{\sqrt{\tau_\theta}}[\lambda(1-\delta) - \delta(\lambda - \omega_t)] < 0 \qquad \forall \omega_t \tag{11}$$

meaning that from $-\infty$ *to* $+\infty$, *Var[*$w_t^*|w_t^* > w_t^R$*] decreases monotonically from 1 to 0 by continuity.*

Thus, over time, the only people who stay with the same employer must have drawn from the upper portion of the distribution.

When jobs are pure experience goods, which is the case where $\tau_m \to 0$, the quality of the match is revealed through job experience. Workers know that they can obtain a starting wage equal to μ for any draw made from the prior distribution. In contrast to the previous case, this is the only outside offer w_a they will receive over time from the outside as nothing is revealed about the quality of a new match short of working in it. Therefore, the decision to quit hinges completely on the belief workers form about the quality of the match. From Eq. (6) above, we know the unconditional distribution of the wage received by a sample of workers. This would be the eventual wage distribution if all workers were to stay with the same employer for their entire career. However, as described in the workers' optimal decision rule above, some workers will quit as the information they get from the sequence of output indicates with more and more precision that it may be preferable to accept a wage equal to μ in a new job, which offers better upward potential. Consequently, the moments of the truncated distribution are similar to those given in Proposition 2, replacing w_t^*, and using σ_t^2=Var[w_t] as given in Eq. (6). I summarize these results in the following proposition.

Proposition 4. *When jobs are pure experience goods, i.e. as* $\tau_m \to 0$, *such that an employee works at a job that pays* $w_t^* > w_t^R$ *for an optimally determined* w_t^R, *then the first two moments of the truncated distribution of his wages are given by*

$$E[w_t^*|w_t^* > w_t^{\mathrm{R}}] = \mu + \sigma_t \frac{\phi\left[\sqrt{\tau_\theta}(w_t^{\mathrm{R}} - \gamma)\right]}{1 - \Phi\left[\sqrt{\tau_\theta}(w_t^{\mathrm{R}} - \gamma)\right]} = \mu + \sigma_t \lambda(\omega_t) \tag{12}$$

$$\mathrm{Var}[w_t^*|w_t^* > w_t^{\mathrm{R}}] = \sigma_t^2\left[1 + \omega_t\lambda(\omega_t) - (\lambda(\omega_t))^2\right] = \sigma_t^2[1 - \delta(\omega_t)] \tag{13}$$

As in the case where jobs are pure search goods, the derivative of V could be calculated to get the evolution of the variance within jobs. Given that σ_t is not constant, additional terms appear in the analog of Eq. (11), hence the effect of the previous wage is in general indeterminate. However, the model *does* imply that from $t=0$ to $t=1$, there is an increase in the variance, as $\sigma_0 = 0$ and $\sigma_1 > 0$. After that, it may increase or decrease depending upon the speed at which the learning process evolves. If all learning is done rapidly, then a selection process identical to the one in case above would dominate and impart a decreasing pattern to the variance within jobs.

To recapitulate, if jobs are pure search or inspection goods, the variance of wages for job stayers will decline. If jobs are pure experience goods, the variance is characterized by an increase from $t=0$ to $t=1$. For $t>1$, the evolution is undetermined. If the learning process is completed fairly rapidly, then we would expect the selection process to impart a decreasing pattern like in the case where jobs are pure search goods. Naturally, one would expect that most jobs would be a mixture of these two extreme cases. As documented by Farber (1994), the weekly hazard out of employment peaks at around 3 months. Such a pattern is not consistent with a pure heterogeneity in the hazards explanation, but is consistent with some learning taking place early on. Thus, it appears that most jobs have to be experienced at least for a while before workers know all they need to know about those jobs.

3. The theory of human capital and the covariance structure

The human capital earnings function has been used extensively since its development by Mincer (1974). However, the implications in terms of the covariance structure have rarely been studied, with Hause (1980) and Kearl (1988) being notable exceptions. This section will first summarize these implications before turning to the empirical section.

Mincer's overtaking concept is at the heart of the analysis in terms of the covariance structure of earnings. Assuming that two individuals have identical attributes (total labour market experience, schooling, etc.), then, according to human capital theory, their earnings profiles should differ only if their rates of post-schooling investment are not the same. The worker who invests more should have lower initial earnings than the other worker. As she accumulates human capital, her earnings should increase at a faster rate and eventually surpass the other worker's earnings. Generalizing to a sample of relatively homogeneous workers followed over time, we should then observe a declining profile of the variance of the log of earnings in the years leading to the overtaking point, followed by an increase in variance afterwards. Therefore, we should have a negative correlation between the slope of the earnings profile and its intercept.

Note that the observed time profile of variances might not necessarily be U-shaped where individuals differ in some unobserved dimension that is correlated with the rate of investment, as Mincer points out in his analysis of residuals.[6] Assuming that high-ability individuals invest more in on-the-job training, then we might have a *positive* correlation between the slope and the intercept of the earnings profile if the correlation between unobserved "ability" and the rate of investment is sufficiently strong.

These empirical implications of the theory of human capital have been tested by Hause (1980) with a 6-year sample of young Swedish white collars who were born the same year and had similar levels of schooling. He allows each individual to have his own level and slope (i.e. experience slope) parameters, implying a residual structure of earnings of the form:[7]

$$w_{it} = \alpha_i + \beta_i X_{it} + u_{it} \tag{14}$$

where X_{it} is individual i's experience at time t and with $E[u_{it}] = 0$ and $\mathrm{Var}[u_{it}] = \sigma_{ut}^2$. The slope and intercept parameters are assumed to be independent of the residual term and also

[6] For the full derivation of these assertions, see Mincer (1974), chapter 6.

[7] Actually, Hause directly fit the wage observations instead of using the residuals from a fully specified earnings function.

independently distributed across individuals. Consequently, taking the expectation over individuals of the cross-product of u_{it} and u_{is} gives us:

$$E[w_{it}w_{is}] = \phi_{\alpha} + \phi_{\beta}X_{it}X_{is} + \phi_{\alpha\beta}(X_{it} + X_{is}) + \sigma_{ut}^{2} \tag{15}$$

$\phi_{\alpha} = \mathrm{Var}[\alpha_i]$, $\phi_{\beta} = \mathrm{Var}[\beta_i]$ and $\phi_{\alpha\beta} = \mathrm{Cov}[\alpha_i, \beta_i]$. Relating Eq. (15) to the discussion above, the overtaking year should be associated with the year at which we observe the minimum variance. Minimizing Eq. (15) with respect to the level of experience X_{OT} (i.e. time of overtaking) implies $X_{\mathrm{OT}} = -\phi_{\alpha\beta}/\phi_{\beta}$. Due to the fact that unobserved ability tends to cancel the predicted trade-off, X_{OT} will represent a lower bound of the true overtaking year. With 6 years of data, the empirical covariance matrix contains 21 distinct elements. Estimating this matrix and incorporating the structure implied by the theory, Hause finds a negative correlation between the slope and intercept parameters. However, whether this correlation is significant or not depends on the residual error structure. Under the specification that produces a statistically significant result, X_{OT} is estimated to be in the neighborhood of 5 years. It should be mentioned that the time profile of the variances of the log of earnings is not U-shaped in Hause's data. Instead, the variances show a monotonic decrease from year 1 to year 6. This is one pattern that learning theories, like the theory of matching, predict, especially in the formulation where all learning about the job characteristics is done rapidly. In fact, when Hause allows the u_{it} term to be independently (but not identically) distributed, he does not get a significant trade-off between the slope and the intercept. This is not surprising since the six distinct residual variance parameters (σ_{ut}^{2}) are picking up the monotonic decrease in the variance of the log of earnings. That leaves essentially no role for the covariance parameter.

Therefore, to discriminate between the two theories, it is not sufficient to study the variance of the experience profiles of workers alone. The next section extends Hause's approach to incorporate tenure profiles.[8]

4. Empirical implementation

4.1. The data

The predictions derived above are examined using unbalanced data from the National Longitudinal Survey of Youth (NLSY) for the period spanning the years 1979–1996. This data set contains the full employment history of young Americans from the moment they make their first long-term transition to the labor force. I essentially use the same sample selection criteria as those used by Farber and Gibbons (1996). Namely, I classify individuals as having made a long-term transition to the labor force when they spend at

[8] See also Light and McGarry (1998) for related work in which they show, using wage models, that both the level and the tenure slope of the log-wage profile are affected by the pattern of job mobility. They conclude that their results are consistent with experience good models of job mobility. The results in this paper suggest that while matching effects are indeed important, they seem to be played out very quickly.

Table 1
Mean sample characteristics (weighted) — NLSY

	Men	Women
Real hourly wage ($1979)	6.20	5.18
Tenure	3.19	3.20
Experience	6.50	5.96
Years in school	13.41	13.63
Percentage nonwhite	19.83	19.85
Percentage married	45.71	45.61
Age	27.25	27.12
Percentage covered by CBA	17.64	15.36
Number of observations	24,799	22,453
Number of individuals	2990	2915
Number of jobs	9583	7910

least three consecutive years primarily working, following a year spent primarily not working.

Someone is classified as primarily working if she/he has worked at least half the weeks since the last interview and averaged at least 30 h/week during the working weeks. See Farber and Gibbons (1996) for more details. Self-employed workers are deleted, as are members of the NLSY military subsample. After some experimentation, I chose to use the full NLSY sample, instead of using only the original representative cross section. As is well known, the NLSY includes a supplemental sample designed to oversample civilian Hispanic, black, and economically disadvantaged nonblack, non-Hispanic youth. The results are very similar across samples (but not across gender), and thus I want to benefit from the largest possible sample size so as to improve precision in the estimates. I am then left with a sample of 47,252 wage observations on 5905 workers. Since all models will be estimated separately for men and women, summary statistics for both are reported in Table 1.

4.2. Log-earnings equation

Let the wage of person i at time t be determined according to the following equation:

$$w_{ijt} = E\left[w_{ijt} \mid \text{observables}\right] + \epsilon_{ijt} \tag{16}$$

where w_{ijt} is the log of the real hourly wage of worker i in job j at time t and ϵ_{ijt} has mean 0. The first step is then to estimate Eq. (16) with ordinary least squares to obtain the residuals. After that, the models that I estimate are models for the expectation of the cross-products of the residuals, $E[\epsilon_{ijt}\epsilon_{iks}]$.

The vector of observables includes years of completed schooling, experience, experience squared, tenure, tenure squared, experience in the industry, experience in the industry squared, industry, occupation, year, race, marital status, union membership, and residence in an SMSA dummies, as well as the local unemployment rate in the county of residence. The vector of estimated residuals is then expressed in deviations from annual means.

4.3. Econometric models of matching

As a first step, a test of the matching model in its simplest form is presented. More precisely, I test whether the covariance structure of the error term of a standard human capital earnings equation satisfies the restrictions imposed by the job matching theory. The simplest way to account for the process of matching is to include an error component θ_{ij} as part of the total error term ϵ_{ijt}. It represents the unobserved (to the econometrician) quality of the match that affects the wage of individual *i* in job *j*. This component is assumed to be fixed within matches although there is a whole distribution of match productivities. Assuming that the error term also contains an individual-specific component to reflect time-invariant unobserved individual characteristics, then it can be written as:

$$\epsilon_{ijt} = \alpha_i + \theta_{ij} + \eta_{ijt} \tag{17}$$

where η_{ijt} is a white noise error term. Note that the three terms are assumed to be independently distributed. Taking the expectation of the cross-products gives:

Model 1

$$E\left[\epsilon_{ijt}^2\right] = \sigma_\alpha^2 + \sigma_\theta^2 + \sigma_\eta^2 \tag{18}$$

$$E[\epsilon_{ijt}\,\epsilon_{iks}] = \sigma_\alpha^2 + I_i(j=k)\sigma_\theta^2 \tag{19}$$

where $I_i(j=k)$ is an indicator variable equal to 1 if worker *i* holds the same job at periods *t* and *s*.

Flinn (1986) provides a statistical test of these restrictions. Using data from the National Longitudinal Survey of Young Men (248 individuals followed from 1967 to 1969), he does not reject the restrictions implied by the theory. An important first step in this paper is thus to replicate Flinn's analysis by using the much larger and longer NLSY.

To test the predictions of the theory of matching pertaining to the evolution of the variance of the job-match component within jobs, model 1 is generalized with the use of dummy variables for different levels of tenure. Let dumten_{il} be equal to 1 if the tenure level of worker *i* is included in the interval corresponding to *l* and equal to 0 otherwise, where $l = \{1$ if tenure < 1 month, 2 if 1 month < tenure < 4 months, 3 if 4 months < tenure < 1 year, 4 if 1 year < tenure < 2 years,. . .,15 if tenure > 11 years). The inclusion of dummies for low levels of tenure captures the increase (if there is any) in the variance of the job-match component, a prediction of the matching model where at least some learning about the job occurs in the course of the employment relationship. If it is impossible to isolate an increasing pattern with this specification, then it would suggest that most of the on-the-job learning process is completed very quickly. The model for the within-job evolution is then:

Model 2

$$E\left[\epsilon_{ijt}^2|\text{dumten}_{il}\right] = \sigma_\alpha^2 + \sum_{l=1}^{15}(\text{dumten}_{il}\sigma_{\theta l}^2) + \sigma_\eta^2 \tag{20}$$

$$E\left[\epsilon_{ijt}\epsilon_{iks}|I_i,\ \text{dumten}_{il}\right] = \sigma_\alpha^2 + I_i(j=k)\sum_{l=1}^{15}(\text{dumten}_{il}\sigma_{\theta l}^2) \tag{21}$$

4.4. Econometric model of human capital

Generalizing Hause's model to the tenure profile, let the error term be specified as:

$$\epsilon_{ijt} = \alpha_i + \theta_{ij} + \beta_i X_{it} + \gamma_i T_{ijt} + \eta_{ijt} \tag{22}$$

where T_{ijt} is i's tenure on job j, at time t. The left-hand side residual is estimated from Eq. (16), which contains quadratic functions of both tenure and experience. The inclusion of these two explanatory variables in the error structure implies a random coefficient model, which is consistent with Mincer's derivation of the human capital function.

Mincer (1974, chapter 5, Section 5.2) includes a subscript i representing each individual with the idea that the rate of investment as well as the return on that investment are individual-specific parameters.

"If information were available on all variables and parameters for each individual i, the [earnings] equation would represent a complete accounting (...) of the human capital characteristics entering into the formation of earnings" (Mincer, 1974, p. 90). Since these parameters are not directly observable, their corresponding i subscript is then eliminated and we are left with estimating average rates of returns. By focusing on the second moments, it is possible to take into account the randomness of the tenure and experience slope parameters. Thus, the model to be estimated has the following basic structure:

Model 3

$$w_{it} = b_X X_{it} + b_T T_{ijt} + \epsilon_{ijt} \tag{23}$$

$$\epsilon_{ijt} = \alpha_i + \theta_{ij} + \beta_i X_{it} + \gamma_i T_{ijt} + \eta_{ijt} \tag{24}$$

$$E[\epsilon_{ijt}] = E[\alpha_i] = E[\theta_{ij}] = E[\beta_i] = E[\gamma_i] = E[\eta_{ijt}] = 0 \tag{25}$$

$$E\left[\epsilon_{ijt}^2 \mid X_{it}, T_{ijt}\right] = \sigma_\alpha^2 + \sigma_\theta^2 + \sigma_\beta^2 X_{it}^2 + \sigma_\gamma^2 T_{ijt}^2 + \sigma_{\alpha\beta} 2X_{it} + \sigma_{\theta\gamma} 2T_{ijt} + \sigma_\eta^2 \tag{26}$$

$$\begin{aligned} E\left[\epsilon_{ijt}^2 \mid X_{it}, T_{ijt}, X_{is}, T_{iks}, I_i\right] = {} & \sigma_\alpha^2 + \sigma_\beta^2 X_{it} X_{is} + \sigma_{\alpha\beta}(X_{it} + X_{is}) + I_i(j = k) \\ & \times \left[\sigma_\theta^2 + \sigma_\gamma^2 T_{ijt} T_{iks} + \sigma_{\theta\gamma}(T_{ijt} + T_{iks})\right] \end{aligned} \tag{27}$$

where T and X are respectively the tenure and experience levels and I_i is the same indicator variable as in the previous section.[9] Note the following assumptions: (a) α_i, θ_{ij} and η_{ijt} are

[9] Other covariates, including experience squared and tenure squared, are not shown in the wage equation although they are included in the estimation.

independent from one another, as previously, (b) β_i and γ_i are also independently distributed, and (c) experience X_{it} and tenure T_{ijt} are not correlated with the random parameters. Therefore, I am assuming away any selection effects that would occur were tenure to be positively correlated with the random coefficient γ_i. This orthogonality assumption is likely to be a strong one as we would expect workers with a high γ_i to stay longer on their job. Also, I am assuming that the observable individual characteristics used in estimating Eq. (26) are independent of the unobservables.[10]

According to human capital theory, we should observe a negative correlation between job-specific slope and intercept parameters. In other words, workers who have relatively more on-the-job training should pay for it through lower initial earnings. Note that the same counter-effect that made Hause's estimate of Mincer's overtaking point a lower bound is present here. More precisely, Jovanovic (1979a) shows that the initial quality of the match and the level of investment in firm-specific capital are complementary. In other words, this element of selection into firm-specific training episodes based on the unobserved quality of the match should create a *positive* correlation between the job-specific slope and intercept parameters, not a negative one. Thus, the U-shaped pattern of the variance of the log of earnings within jobs is not a prediction that is shared by the two theories.

Another prediction of human capital theory is that training should occur early in one's career (e.g. Ben-Porath, 1967). Therefore, in the absence of matching effects, the U-shaped pattern should be more evident in the first few jobs than later on. However, if a worker entering the labour market and a firm find themselves in a bad match, they may be less likely to invest in firm-specific skills than if the match was a good one. If these job-matching considerations are at play, then the correlation between the job-specific intercept and slope parameters should tend to be weaker in the first job than in subsequent jobs. To verify these predictions, Eq. (26) is expanded by fitting four different quadratic functions to allow examination of the evolution of the parameters from job to job.

One potential problem with trying to fit different functions for each job is that the subsample of workers with four or more jobs may be either quite small and/or not very representative. This is especially true in the case of workers with at least some college. Given the age structure of the NLSY in 1979, where the age of the respondents ranged from 14 to 21, I may not have such a large number of more educated workers who go through at least four jobs. To gauge whether this might be a real problem, Table 2 displays the sample size by educational attainment and tenure levels for each job.[11] We can see that once we reach job 4+, the number of observations at higher levels of tenure, say over 6 years, is quite small. This is basically true irrespective of the education level.

[10] Ideally, one would like to allow for arbitrary correlations between the observables and the variance components as well as between the parameters and their associated observables. However, one very rapidly runs into identification problems if such a strategy is adopted. Again, I should stress that one of the interesting features of the approach followed in this paper is that matching cannot predict a trade-off between slope and intercept. Consequently, even if the random parameters are correlated with the variables, it is not clear how that would affect the overall conclusions.

[11] Note that the same individual can appear in more than one education category if, for example, that individual first transited "permanently" only to return to school for a few years and then transit to the labor force again. Also, individuals' educational attainment can change even when one working full-time.

Table 2
Number of wage observations per level of education and tenure for each job

Education and employer tenure	Job 1	Job 2	Job 3	Jobs 4+
Panel A: Men				
High school dropouts				
0–2 years	728	516	352	571
2–4 years	305	163	108	131
4–6 years	157	71	50	48
6–8 years	86	44	26	15
8–10 years	61	30	17	4
10 + years	81	30	9	6
High school graduates				
0–2 years	1646	1475	1004	1439
2–4 years	944	643	364	468
4–6 years	562	384	193	180
6–8 years	360	227	80	91
8–10 years	223	139	44	40
10 + years	321	112	26	15
At least some college				
0–2 years	1542	1351	859	1030
2–4 years	969	647	381	354
4–6 years	630	398	199	147
6–8 years	414	242	89	61
8–10 years	278	124	47	12
10 + years	299	91	38	8
Panel B: Women				
High school dropouts				
0–2 years	443	234	121	107
2–4 years	254	66	42	32
4–6 years	114	39	20	12
6–8 years	63	14	5	8
8–10 years	37	2	5	3
10 + years	34	1	1	2
High school graduates				
0–2 years	1635	1220	721	700
2–4 years	1051	506	276	241
4–6 years	628	282	133	92
6–8 years	379	150	76	38
8–10 years	248	92	25	16
10 + years	329	62	30	13
At least some college				
0–2 years	1908	1630	1043	958
2–4 years	1178	796	454	347
4–6 years	740	426	231	140
6–8 years	451	263	124	63
8–10 years	321	143	66	26
10 + years	392	102	39	10

In part to evaluate when the predicted trade-off occurs, and also to put more weight on the relatively low-tenure observations, I also expanded the above model to have different quadratic functions for each job and for different tenure intervals. After some experimentation, I settled on four levels of tenure (less than a year, 1–3 years, 3–5 years, 5+ years) for jobs 1 to 4+. One reason why this exercise might be of interest is that it may allow us to observe whether job-match gains following a job change translate into firm-specific training occurring perhaps earlier during the employment relationship. Indeed, using data from the training module of the NLSY, Loewenstein and Spletzer (1997) found that much (largely formal) training seemed to be delayed instead of occurring up-front, as the standard theory of human capital would suggest. They interpreted those results as evidence of imperfect information about the quality of a job match, which would tend to delay the investment decision.

Finally, given the well-known concave wage-experience profile, I fit two different functions for the predicted trade-off between the individual-specific intercept and the individual-specific experience slope, to capture the fact that most of the wage growth occurs early. One was estimated for observations with 8 years of experience of less, and another was estimated for higher levels of experience.[12]

4.5. Estimation methodology

To estimate and test the restrictions imposed by the models, I make use of the methodology proposed by Gallant and Jorgenson (1979) in the context of a system of nonlinear implicit equations. The basic idea is to compare the weighted sum of squares of an unrestricted model with that of the (restrictive) model I wish to estimate. This methodology is closely related to the minimum distance method proposed by Chamberlain (1984) and adapted by Abowd and Card (1989) to study the covariance structure of earnings and hours changes. The conditional moments equations are stacked up into a system of equations, which, as an example, would be of the following form for model 1:

$$E[m_i \mid I_i] = f(\beta, I_i) \tag{28}$$

$$m_i = \left[\epsilon_{i1}^2, \epsilon_{i1}\epsilon_{i2}, \epsilon_{i1}\epsilon_{iT_i}, \epsilon_{i2}^2, \epsilon_{i2}\epsilon_{i3}, \ldots \epsilon_{iT_i}^2\right] \tag{29}$$

where T_i is the last period in which worker i is in the sample, β is the vector of parameters to estimate and $f(\cdot,\cdot)$ is the mapping representing the model. Were all 5905 individuals present in the sample from 1979 to 1996,[13] each model for a given worker

[12] I could have added squared and cubic terms in both experience and tenure to Eq. (24) to accommodate nonlinearities. However, this would translate into a very large number of additional parameters to estimate once all cross-products would be computed.

[13] Note that no interview took place in 1995.

would represent a system of 153 equations (17 contributions to the variances and 136 contributions to the covariances), for a grand total of 903,465 observations of cross-products. The sample being unbalanced, I end up with 131,627 observations for 2990 males and 113,722 observations for 2915 females. The objective is then to minimize the following function:

$$S(\beta) = \sum_{i=1}^{N} \left[m_i - f(\beta,I_i)\right]' V^{-1} \left[m_i - f(\beta,I_i)\right] \tag{30}$$

Where V^{-1} is computed with the cross-products of the residuals from the following unrestricted model:

$$E[\epsilon_{it}^2] = \text{cov}1_{tt} \tag{31}$$

$$E\left[\epsilon_{it}\epsilon_{is}|I_i\right] = \text{cov}1_{ts} + I_i\text{cov}2_{ts} \tag{32}$$

This unrestricted model contains 153 different cov1 parameters and 136 distinct cov2 parameters. Note that if only the cov1s were estimated, each of these parameters would be equal to the corresponding sample moment. Nesting the matching models in a more general model that depends on each individual-specific I_i requires the estimation of the cov2s as well. Let $S(\beta)_U$ be the value of the objective function for the unrestricted model and $S(\beta)_R$ be defined likewise for the restricted version of Eq. (7). For more general (nonlinear) models, Gallant and Jorgenson (1979) show that $T_0 = S(\beta)_R - S(\beta)_U$ is asymptotically distributed as a chi-square with $r-s$ degrees of freedom where $r-s$ is equal to the difference in the number of parameters in the two models.

5. Results

The results from estimating the simple component of variance model outlined in Eq. (17) for the entire samples of men and women are reported in Table 3. I first report the results from estimating model 1 without the match component ("model 0"). Not too surprisingly, I find little statistical support for the model that provides control for individual heterogeneity only and in which t is i.i.d. The distance statistics of 3867 for men and 3161 for women are surprising values coming from a χ^2 (287).[14] Even when the job-match component is added, the structure is still decidedly rejected, unlike the results in Flinn (1986). There is nevertheless a substantial improvement in the fit of the model.

Looking at Table 4 and the within-job evolution of the variance of the job-match component (model 2) by education level, we can see that the general pattern is one of an overall decline occurring rather early. Only in the case of female high school graduates do we see an increase in the contribution of the variance component during the first year.

[14] In fact, the p-values are essentially 0 for all models estimated in Table 3.

Table 3
Contribution of job-match component (standard errors in parentheses)

Parameter	Model 0	Model 1
Panel A: Men		
Variance of unmeasured and fixed worker ability, ${\sigma_\alpha}^2$	0.0258 (0.0018)	0.0160 (0.0017)
Variance of unmeasured and fixed job-match quality, ${\sigma_\theta}^2$	–	0.0277 (0.0003)
Variance of residual white noise error term, ${\sigma_\eta}^2$	0.0427 (0.0013)	0.0333 (0.0020)
χ^2 statistic (degrees of freedom)	3867 (287)	1001 (286)
Number of workers	2990	2990
Number of cross-products	130,627	130,627
Panel B: Women		
Variance of unmeasured and fixed worker ability, ${\sigma_\alpha}^2$	0.0347 (0.0018)	0.0245 (0.0018)
Variance of unmeasured fixed job-match quality, ${\sigma_\theta}^2$	–	0.0210 (0.0003)
Variance of residual white noise error term, ${\sigma_\eta}^2$	0.0290 (0.0011)	0.0219 (0.0011)
χ^2 statistic (degrees of freedom)	3161 (287)	703 (286)
Number of workers	2915	2915
Number of cross-products	113,722	113,722

While the variance of the job-match component at around 2 years of tenure tends to be smaller than initially (the exception being female high school graduates), the only reasonably solid conclusion one can draw from these results is that matching effects seem to occur very early in the employment relationships and are rapidly played out, as the results in Farber (1994) strongly suggested. This result is also similar to what Topel and Ward (1992) found using Social Security earnings records.

I turn next to Table 5 for the basic extension of the Hause model.[15] Note that to accommodate the fact that the workers in my samples have accumulated as many as 18 years of labour market experience, I fit two linear trends in experience. For comparison purposes with Hause (1980), I report the results when only two linear trends are estimated (see Eq. (14)), both with and without the term representing the variance of the job-match component. The predicted trade-off between the experience slope and the individual intercept for those with 8 years or less of experience ($\sigma_{\alpha_1},\beta_1$) comes out quite strongly when ${\sigma_\theta}^2$ is absent from the equation. However, once I include it, the estimated covariance is smaller by about 30%. Turning now to the estimation of model 3, we can see that both for the first linear experience trend and for the linear tenure trend, the covariance term is estimated quite precisely. Also, the magnitude of the coefficients is larger for men than it is for women, both for the experience profile and the tenure profile. It also noteworthy to observe that there is no evidence of the predicted trade-off for the experience profile past

[15] Note that I have not computed χ^2 test statistics for these models, which include functions of experience and tenure as regressors. While a relatively simple weighing scheme as in Farber and Gibbons (1996) can be used to test the fit of the simple component of variance models that take into account either unobserved worker ability and unobserved match quality, it is not clear how one can do so with regressors other than dummy variables when dealing with unbalanced data.

Table 4
Contribution of job-match parameter by tenure level (standard errors in parentheses)

Parameter	Less than high school	High school degree	At least some college
Panel A: Men			
Variance of unmeasured match quality, σ_θ^2			
Tenure < 4 months	0.0382 (0.0068)	0.0530 (0.0049)	0.0904 (0.0096)
Tenure ∈ [4 months, 1 year)	0.0330 (0.0049)	0.0507 (0.0033)	0.0614 (0.0062)
Tenure ∈ [1–2 years)	0.0398 (0.0041)	0.0354 (0.0027)	0.0559 (0.0045)
Tenure ∈ [2–3 years)	0.0274 (0.0033)	0.0266 (0.0022)	0.0433 (0.0033)
Tenure ∈ [3–4 years)	0.0308 (0.0039)	0.0263 (0.0025)	0.00290 (0.0038)
Tenure ∈ [4–5 years)	0.0180 (0.0047)	0.0238 (0.0030)	0.0316 (0.0043)
Tenure ∈ [5–6 years)	0.0251 (0.0054)	0.0109 (0.0033)	0.0531 (0.0049)
Tenure ∈ [6–7 years)	0.0229 (0.0074)	0.0153 (0.0039)	0.0231 (0.0055)
Tenure ∈ [7–8 years)	0.0427 (0.0089)	0.0053 (0.0043)	0.0249 (0.0062)
Tenure ∈ [8–9 years)	0.0542 (0.0104)	0.0172 (0.0053)	0.0117 (0.0072)
Tenure ∈ [9–10 years)	0.0346 (0.0113)	0.0098 (0.0060)	0.0055 (0.0088)
Tenure ∈ [10–11 years)	0.0263 (0.0124)	0.0256 (0.0066)	0.0165 (0.0098)
Tenure ∈ [11–12 years)	0.0570 (0.0150)	0.0170 (0.0086)	0.0166 (0.00131)
Tenure ∈ [12–13 years)	−0.0149 (0.0169)	0.0165 (0.0085)	0.0029 (0.0136)
Tenure ≥ 13 years	−0.0099 (0.0144)	0.0001 (0.0070)	−0.0151 (0.0113)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	567	1437	1334
Number of cross-products	17,325	57,710	50,953
Panel B: Women			
Variance of unmeasured match quality, σ_θ^2			
Tenure < 4 months	0.0507 (0.0163)	0.0255 (0.0075)	0.0683 (0.0077)
Tenure ∈ [4 months, 1 year)	0.0345 (0.0115)	0.0489 (0.0048)	0.0436 (0.0045)
Tenure ∈ [1–2 years)	0.0325 (0.0099)	0.0451 (0.0035)	0.0437 (0.0035)
Tenure ∈ [2–3 years)	0.0285 (0.0077)	0.0257 (0.0026)	0.0272 (0.0025)
Tenure ∈ [3–4 years)	0.0217 (0.0074)	0.0283 (0.0029)	0.0313 (0.0027)
Tenure ∈ [4–5 years)	0.0207 (0.0073)	0.0261 (0.0034)	0.0407 (0.0032)
Tenure ∈ [5–6 years)	0.0196 (0.0077)	0.0208 (0.0039)	0.0251 (0.0037)
Tenure ∈ [6–7 years)	0.0055 (0.0096)	0.0228 (0.0042)	0.0271 (0.0043)
Tenure ∈ [7–8 years)	−0.0109 (0.0100)	0.0208 (0.0050)	0.0309 (0.0050)
Tenure ∈ [8–9 years)	0.0365 (0.0116)	0.0367 (0.0059)	0.0465 (0.0057)
Tenure ∈ [9–10 years)	0.0294 (0.0130)	0.0197 (0.0069)	0.0193 (0.0071)
Tenure ∈ [10–11 years)	0.0193 (0.0152)	0.0139 (0.0074)	0.0192 (0.0072)
Tenure ∈ [11–12 years)	0.0064 (0.0173)	0.0081 (0.0088)	0.0262 (0.0089)
Tenure ∈ [12–13 years)	0.0104 (0.0230)	0.0101 (0.0096)	0.0112 (0.0094)
Tenure ≥ 13 years	−0.0218 (0.0177)	0.0036 (0.0090)	−0.0023 (0.0085)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	313	1296	1611
Number of cross-products	6,821	44,004	58,288

the first 8 years of labour market experience. In some sense, it should not be all that surprising given the well-known flattening of experience-wage profile after the first few years of rapid growth.

Table 5
Covariance structure of residuals as function of tenure, experience, unobserved components (entire sample; standard errors in parentheses)

Parameter	Hause's model without match component	Hause's model with match component	Model 3+
Panel A: Men			
First 8 years of experience			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_1}^2$	0.0939 (0.0039)	0.0618 (0.0040)	0.0583 (0.0040)
Variance of experience slope, $\sigma_{\beta_1}^2$	0.0018 (0.0001)	0.0013 (0.0001)	0.0013 (0.0001)
Covariance of experience and slope worker-specific intercept, $\sigma_{\alpha_1,\beta_1}$	−0.0044 (0.0005)	−0.0030 (0.0005)	−0.0025 (0.0005)
Experience>8 years			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_2}^2$	0.0787 (0.0112)	0.0525 (0.0112)	0.0490 (0.0112)
Variance of experience slope, $\sigma_{\beta_2}^2$	0.0002 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Covariance of experience and slope worker-specific intercept, $\sigma_{\alpha_2,\beta_2}$	0.0008 (0.0011)	0.0008 (0.0011)	0.0013 (0.0011)
Variance of unobserved match quality (job-specific intercept), σ_{θ}^2	–	0.0337 (0.0011)	0.0436 (0.0016)
Variance of tenure slope, σ_{γ}^2	–	–	0.0003 (0.0001)
Covariance of tenure slope and job-specific intercept, $\sigma_{\gamma,\theta}$	–	–	−0.0028 (0.0005)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	2990	2990	2990
Number of cross-products	131,627	131,627	131,627
Panel B: Women			
First 8 years of experience			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_1}^2$	0.0750 (0.0034)	0.0441 (0.0036)	0.0419 (0.0036)
Variance of experience slope, $\sigma_{\beta_1}^2$	0.0015 (0.0001)	0.0010 (0.0001)	0.0011 (0.0001)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_1,\beta_1}$	−0.0028 (0.0005)	−0.0016 (0.0005)	−0.0013 (0.0005)
Experience>8 years			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_2}^2$	0.0675 (0.0123)	0.0453 (0.0122)	0.0423 (0.0122)
Variance of experience slope, $\sigma_{\beta_2}^2$	0.0002 (0.0001)	0.0002 (0.0001	0.0002 (0.0001)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_2,\beta_2}$	0.0010 (0.0012)	0.0006 (0.0012)	0.0009 (0.0012)
Variance of unobserved match quality (job-specific intercept), σ_{θ}^2	–	0.0319 (0.0012)	0.0368 (0.0017)
Variance of tenure slope, σ_{γ}^2	–	–	0.0001 (0.0000)
Covariance of tenure slope and job-specific intercept, $\sigma_{\gamma,\theta}$	–	–	−0.0011 (0.0004)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	2915	2915	2915
Number of cross-products	113,722	113,722	113,722

Table 6
Covariance structure of residuals as function of tenure, experience, unobserved components (by education level; standard errors in parentheses)

Parameter	Less than high school	High school degree	At least some college
Panel A: Men			
First 8 years of experience			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_1}^2$	0.0049 (0.0052)	0.0512 (0.0047)	0.0513 (0.0068)
Variance of experience slope, $\sigma_{\beta_1}^2$	0.0007 (0.0002)	0.0011 (0.0001)	0.0012 (0.0001)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_1,\beta_1}$	0.0009 (0.0007)	−0.0024 (0.0005)	−0.0011 (0.0008)
Experience>8 years			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_2}^2$	0.0423 (0.0209)	0.0417 (0.0128)	0.0701 (0.0198)
Variance of experience slope, $\sigma_{\beta_2}^2$	0.0005 (0.0002)	0.0001 (0.0001)	0.0005 (0.0002)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_2,\beta_2}$	−0.0023 (0.0020)	0.0011 (0.0012)	−0.0006 (0.0019)
Variance of unobserved match quality (job-specific intercept), σ_{θ}^2	0.0341 (0.0030)	0.0368 (0.0018)	0.0508 (0.0029)
Variance of tenure slope, σ_{γ}^2	0.0004 (0.0001)	0.0004 (0.0001)	0.0001 (0.0001)
Covariance of tenure slope and job-specific intercept, $\sigma_{\gamma,\theta}$	−0.0017 (0.0007)	−0.0031 (0.0004)	−0.0030 (0.0006)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	567	1437	1334
Number of cross-products	17,325	57,710	50,953
Panel B: Women			
First 8 years of experience			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_1}^2$	0.0045 (0.0110)	0.0158 (0.0043)	0.0514 (0.0049)
Variance of experience slope, $\sigma_{\beta_1}^2$	−0.0006 (0.0006)	0.0010 (0.0001)	0.0013 (0.0001)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_1,\beta_1}$	0.0047 (0.0024)	−0.0007 (0.0006)	−0.0022 (0.0006)
Experience>8 years			
Variance of unmeasured fixed worker ability, $\sigma_{\alpha_2}^2$	0.0641 (0.0392)	0.0307 (0.0160)	0.0331 (0.0170)
Variance of experience slope, $\sigma_{\beta_2}^2$	0.0005 (0.0004)	0.0004 (0.0002)	0.0001 (0.0002)
Covariance of experience slope and worker-specific intercept, $\sigma_{\alpha_2,\beta_2}$	−0.0041 (0.0039)	−0.0005 (0.0016)	0.0025 (0.0017)
Variance of unobserved match quality (job-specific intercept), σ_{θ}^2	0.0311 (0.0077)	0.0385 (0.0024)	0.0349 (0.0021)
Variance of tenure slope, σ_{γ}^2	0.0002 (0.0002)	0.0003 (0.0001)	−0.0001 (0.0001)
Covariance of tenure slope and job-specific intercept, $\sigma_{\gamma,\theta}$	−0.0022 (0.0012)	−0.0027 (0.0005)	−0.0002 (0.0005)
Residual term, 17 i.d. terms	Yes	Yes	Yes
Number of workers	313	1296	1611
Number of cross-products	6821	44,004	58,288

Table 6 shows the results for the full model by education category. The first thing to note is the substantially lower estimate of the unobserved ability component of variance for dropouts with 8 years or less of experience. It seems that either those individuals are

Table 7
Covariance between match-specific intercept and match-specific slope by job and by education level (standard errors in parentheses)

Parameter	Less than high school	High school degree	At least some college
Panel A: Men			
Job 1			
Variance of match component, $\sigma_\theta{}^2$	0.0353 (0.0052)	0.0367 (0.0037)	0.0689 (0.0058)
Variance of tenure slope, σ_γ^2	0.0006 (0.0002)	−0.0001 (0.0001)	−0.0001 (0.0002)
Covariance, $\sigma_{\gamma,\theta}$	−0.0032 (0.0011)	−0.0008 (0.0007)	−0.0030 (0.0010)
Job 2			
Variance of match component, σ_θ^2	0.0423 (0.0049)	0.0371 (0.0030)	0.0611 (0.0047)
Variance of tenure slope, σ_γ^2	0.0004 (0.0003)	0.0005 (0.0002)	0.0008 (0.0003)
Covariance, $\sigma_{\gamma,\theta}$	−0.0034 (0.0014)	−0.0041 (0.0007)	−0.0066 (0.0012)
Job 3			
Variance of match component, σ_θ^2	0.0309 (0.0058)	0.0467 (0.0036)	0.0501 (0.0056)
Variance of tenure slope, σ_γ^2	0.0016 (0.0005)	0.0006 (0.0003)	0.0007 (0.0004)
Covariance, $\sigma_{\gamma,\theta}$	−0.0009 (0.0019)	−0.0063 (0.0011)	−0.0043 (0.0006)
Job 4+			
Variance of match component, σ_θ^2	0.0250 (0.0052)	0.0389 (0.0033)	0.0438 (0.0057)
Variance of tenure slope, σ_γ^2	0.0002 (0.0005)	0.0004 (0.0003)	0.0006 (0.0005)
Covariance, $\sigma_{\gamma,\theta}$	0.0020 (0.0019)	−0.0056 (0.0012)	−0.0052 (0.0019)
Number of workers	567	1437	1334
Number of cross-products	17,325	57,710	50,953
Panel B: Women			
Job 1			
Variance of match component, σ_θ^2	0.0160 (0.0105)	0.0340 (0.0041)	0.0384 (0.0044)
Variance of tenure slope, σ_γ^2	0.0000 (0.0003)	0.0000 (0.0001)	−0.0003 (0.0002)
Covariance, $\sigma_{\gamma,\theta}$	−0.0003 (0.0017)	−0.0012 (0.0007)	0.0005 (0.0008)
Job 2			
Variance of match component, σ_θ^2	0.0202 (0.0120)	0.0431 (0.0036)	0.0489 (0.0035)
Variance of tenure slope, σ_γ^2	0.0007 (0.0011)	0.0003 (0.0002)	0.0001 (0.0002)
Covariance, $\sigma_{\gamma,\theta}$	−0.0008 (0.0036)	−0.0032 (0.0010)	−0.0027 (0.0009)
Job 3			
Variance of match component, σ_θ^2	0.0598 (0.0148)	0.0394 (0.0048)	0.0243 (0.0042)
Variance of tenure slope, σ_γ^2	0.0002 (0.0012)	0.0004 (0.0003)	0.0001 (0.0003)
Covariance, $\sigma_{\gamma,\theta}$	−0.0073 (0.0045)	−0.0047 (0.0014)	−0.0004 (0.0012)
Job 4+			
Variance of match component, σ_θ^2	0.0493 (0.0165)	0.0481 (0.0056)	0.0319 (0.0043)
Variance of tenure slope, σ_γ^2	0.0007 (0.0011)	0.0026 (0.0004)	0.0001 (0.0002)
Covariance, $\sigma_{\gamma,\theta}$	−0.0015 (0.0046)	−0.0088 (0.0017)	−0.0012 (0.0013)
Number of workers	313	1296	1611
Number of cross-products	6821	44,004	58,288

Other estimated parameters include, in addition to 17 terms for the residual variance, components of variance for: (i) unmeasured individual ability, (ii) the variance of the experience profile, (iii) the covariance of the individual-specific intercept and the experience slope, as in Table 5.

Table 8
Covariance between job-specific slope and job-specific intercept by job and tenure range (standard errors in parentheses)

Parameter: $\sigma_{\gamma,\theta}$	Less than high school	High school degree	At least some college
Panel A: Men			
Job 1			
Tenure<1 year	−0.0152 (0.0299)	0.0210 (0.0173)	0.0220 (0.0223)
1 year≤Tenure<3 years	−0.0061 (0.0089)	−0.0023 (0.0053)	−0.0132 (0.0069)
3 years≤Tenure<5 years	0.0114 (0.0130)	0.0029 (0.0074)	0.0070 (0.0094)
5 years≤Tenure	−0.0016 (0.0031)	0.0004 (0.0017)	−0.0036 (0.0025)
Job 2			
Tenure<1 year	0.0376 (0.0292)	−0.0070 (0.0147)	−0.0708 (0.0251)
1 year≤Tenure<3 years	−0.0103 (0.0096)	0.0010 (0.0050)	−0.0091 (0.0070)
3 years≤Tenure<5 years	0.0421 (0.0149)	−0.0079 (0.0062)	0.0051 (0.0093)
5 years≤Tenure	−0.0013 (0.0046)	0.0012 (0.0022)	−0.0040 (0.0038)
Job 3			
Tenure<1 year	−0.1237 (0.0324)	−0.0434 (0.0190)	0.0049 (0.0225)
1 year≤Tenure<3 years	−0.0225 (0.0147)	0.0077 (0.0064)	−0.0122 (0.0088)
3 years≤Tenure<5 years	−0.0190 (0.0181)	0.0067 (0.0081)	−0.0039 (0.0122)
5 years≤Tenure	0.0189 (0.0076)	−0.0069 (0.0041)	0.0060 (0.0058)
Job 4+			
Tenure<1 year	−0.0539 (0.0252)	−0.1095 (0.0161)	−0.0429 (0.0247)
1 year≤Tenure<3 years	−0.0274 (0.0109)	−0.0013 (0.0060)	−0.0038 (0.0092)
3 years≤Tenure<5 years	−0.0853 (0.0170)	−0.0079 (0.0084)	−0.0577 (0.0138)
5 years≤Tenure	−0.0066 (0.0137)	−0.0032 (0.0051)	−0.0059 (0.0090)
Number of workers	567	1437	1334
Number of cross-products	17,325	57,710	50,953
Panel B: Women			
Job 1			
Tenure<1 year	0.0393 (0.0530)	−0.0068 (0.0181)	0.0045 (0.0168)
1 year≤Tenure<3 years	0.0191 (0.0190)	−0.0010 (0.0054)	−0.0039 (0.0052)
3 years≤Tenure<5 years	−0.0091 (0.0216)	−0.0090 (0.0077)	−0.0061 (0.0072)
5 years≤Tenure	0.0052 (0.0040)	−0.0016 (0.0019)	0.0024 (0.0019)
Job 2			
Tenure<1 year	−0.0401 (0.0608)	0.0038 (0.0204)	−0.0273 (0.0199)
1 year≤Tenure<3 years	0.0005 (0.0311)	− 0.0105 (0.0057)	0.0000 (0.0050)
3 years≤Tenure<5 years	0.0144 (0.0423)	−0.0068 (0.0078)	−0.0089 (0.0068)
5 years≤Tenure	−0.0020 (0.0213)	0.0016 (0.0033)	0.0032 (0.0030)
Job 3			
Tenure<1 year	0.2086 (0.1040)	0.0049 (0.0210)	−0.0833 (0.0196)
1 year≤Tenure<3 years	−0.0105 (0.0450)	0.0067 (0.0083)	0.0009 (0.0061)
3 years≤Tenure<5 years	0.0019 (0.0158)	0.0021 (0.0117)	0.0006 (0.0083)
5 years≤Tenure	−0.0034 (0.0259)	−0.0047 (0.0048)	0.0003 (0.0050)
Job 4+			
Tenure<1 year	−0.0183 (0.0935)	−0.0194 (0.0144)	−0.0117 (0.0220)
1 year≤Tenure<3 years	0.0591 (0.0557)	0.0037 (0.0100)	−0.0018 (0.0071)
3 years≤Tenure<5 years	−0.0757 (0.0603)	−0.0169 (0.0136)	0.0118 (0.0113)
5 years≤Tenure	0.0940 (0.0247)	0.0204 (0.0070)	0.0120 (0.0049)

Table 8 (*continued*)

Parameter: $\sigma_{\gamma,\theta}$	Less than high school	High school degree	At least some college
Panel B: Women			
Job 4+			
Number of workers	313	1296	1611
Number of cross-products	6821	44,004	58,288

Other estimated parameters include, in addition to 17 terms for the residual variance, components of variance for: (i) unmeasured individual ability, (ii) the variance of the experience profile, (iii) the covariance of the individual-specific intercept and the experience slope, as in Table 5.

fairly homogeneous to start with or the market pools them to a much greater degree than is the case for more educated workers.[16] This is true for both men and women. In the years following the first few in the labour market, there is a larger dispersion of "types" ($\sigma^2_{\alpha_2}$), although the estimates are rather imprecise.

As for the predicted negative covariance between job-specific slopes and intercepts, it does seem that more education makes a difference for men, but not really for women. For both men and women, education makes a difference for the covariance between the experience slope and the individual intercept (for 8 years or less of experience). We should perhaps be cautious about drawing strong conclusions from the relatively small sample of workers with less than a high school education. Yet, it seems that even dropouts benefit from some investments in firm-specific human capital.

Tables 7 and 8 show the results when the basic model is estimated separately by job or by job and tenure range. The main thing to note about the results in Table 7 is that, in the case of women, only for high school graduates do I find evidence of the predicted trade-off across all jobs. This contrasts with the results for men, particularly the more educated ones. For them, investment seems to take place basically across all jobs. Yet, it is still true that for more educated workers, the covariance parameter is larger in the second job than it is in the first (and significantly so in the statistical sense). As for the results in Table 8, there is some evidence (at least for men) that training opportunities occur earlier in the employment relationship as one moves from the first job to the next ones. This would tend to support the notion that one needs to locate good matches before investments in firm-specific skills take place. For women, there is also certain evidence of this sort of pattern, but the results for them are fairly inconclusive.

5.1. Discussion

Overall, the results strongly suggest that workers who invest more pay for those investments through lower wages. Many researchers (e.g. Barron et al., 1989, 1999) who have contributed to the voluminous literature on the wage impact of training have tried to verify whether this basic prediction of the human capital model received any support from

[16] Years of completed schooling vary through time for some individuals. Thus, the same individuals may appear in more than one education category. That explains why adding the number of workers across columns at the bottom of Table 6 results in a number that is larger than the overall number of individuals in the sample. It also implies that the number of cross-products, when added by education category, will be smaller than in Table 5.

the data. It turns out that it has not. As emphasized by Barron et al. (1989), the confounding effect of unobserved worker ability and the systematic sorting that takes place when workers are allocated to jobs may be a reason why it is so difficult to isolate the impact of training on the initial wage. One could also perhaps invoke the fact that no matter how refined the questions may be on training in a particular survey, it may only provide a fairly rough proxy for the sometimes largely informal nature of training. Consequently, as in all situations where measurement error is a problem, this makes the task of estimating relationships more difficult.

In this paper, I do not use any information on training; instead, I make use of the restrictions implied by the theory on the wage process to test the equivalent prediction that job-specific intercept and job-specific slope parameters should be negatively correlated. In fact, I even find evidence for such a negative correlation in the experience profile. Thus, according to human capital theory, wages should be impacted by whatever form of training is taking place. Looking at the second moments of wages may be the reason I am able to identify some relationships that have proved to be exceedingly difficult to identify when looking only at the first conditional moments.

6. Conclusion

The theories of matching and human capital both predict that wages should be positively correlated with tenure in a cross section of workers. Much of the effort of researchers surrounding the issue of whether this relationship really reflects the accumulation of human capital has focused on trying to break the likely correlation between firm seniority and unobserved worker and job-match characteristics in an earnings function. This paper's objective is to exploit the fact that matching and human capital have different implications for the behavior of the second moments of earnings to assess the relative merits of both theories. Introducing firm-specific wage growth into Hause's (1980) method of testing human capital theory, I am able to find strong evidence suggesting the importance of human capital considerations in the wage formation process. More precisely, the estimated U-shaped pattern of the variance of wage residuals within jobs is a prediction that is exclusive to human capital theory and is not implied by the theory of job matching. Also, there is some evidence that investment in on-the-job training is more likely to occur earlier in the employment relationship as one switches for her/his first job to subsequent jobs.

The selection process implied by matching that predicts a gradual truncation of the wage distribution as workers who draw a good match stay in their jobs longer receives mild support from the data. Given the strong evidence of a negative covariance between job-specific slopes and intercepts, we should probably not be terribly surprised that the sharp prediction of truncation of the wage distribution does not receive stronger support. After all, this prediction is derived under the assumption of no accumulation of human capital.

The main conclusion from the study is that while some of the patterns predicted by the theory of matching are present in the data, it seems difficult to deny the major role played by human capital in the determination of wages.

Acknowledgements

I would like to thank David Albouy, Joe Altonji, David Card, Janet Currie, Boyan Jovanovic, Thomas Lemieux, Bentley MacLeod, Robert Topel, and the seminar participants at Boston College, the Harvard/MIT Labor Seminar, Laval, McMaster, New York University, the Princeton Labor Lunch, McGill, Montréal, and Toronto for their comments and suggestions. Special thanks to co-Editor Dan Black and two anonymous referees for suggestions that helped improve the paper substantially. Part of the revision was done while I was visiting Princeton University. I wish to express my gratitude to this institution for its hospitality. Financial support from the SHRCC is also gratefully acknowledged.

Appendix A

Proof (Proposition 1). Let $f(\ )$ be the posterior distribution, $g(\ )$ the likelihood function, and $h(\ )$ the prior distribution of θ. From Bayes' Law we know that:

$$f(\theta,m,x_1,\ldots,x_t) = g(m,x_1,\ldots,x_t\,|\,\theta)h(\theta) \tag{33}$$

with

$$g(m,x_1,\ldots,x_t\,|\,\theta) \propto \exp\left[-\frac{1}{2}\sum_{1=1}^{t}(x_i-\theta)^2 - \frac{\tau_m}{2}(m-\theta)^2\right] \tag{34}$$

and

$$h(\theta) \propto \exp\left[-\frac{\tau_\theta}{2}(\theta-\mu)^2\right] \tag{35}$$

If we develop Eq. (33) knowing $g(\ )$ and $h(\ )$, then we obtain, after factoring and putting some terms into the multiplicative constant,

$$f(\theta,m,x_1,\ldots,x_t) \propto \exp\left[-\frac{t+\tau_m+\tau_\theta}{2}\left(\theta-\frac{t\bar{x}+\tau_m m+\tau_\theta\mu}{t+\tau_m+\tau_\theta}\right)^2\right] \tag{36}$$

which is the kernel of a normal distribution with conditional expectation given by Eq. (1) and condition variance by Eq. (2).

The useful fact that $S_{t-1}S_t = S_{t-1} - S_t$ can be found be manipulating Eq. (2). Rearranging Eq. (1) and substituting in Eq. (2), we can express w_t as

$$w_t = \frac{S_t}{S_{t-1}}w_{t-1} + S_t[\theta+\epsilon_t] \tag{37}$$

given that $w_{t-1}=w$, and taking the expectation over the posterior distribution we get:

$$E\left[w_t\,|\,w_{t-1}=w\right] = \frac{S_t}{S_{t-1}}w + S_tE\left[\theta+\epsilon_t\,|\,w_{t-1}=w\right] = \frac{S_t}{S_{t-1}}w + S_tw = w \tag{38}$$

using the facts that $E[\theta+\epsilon_t|w_{t-1}=w]=w+0$, and $S_t=1-S_t/S_{t-1}$. Calculating the conditional variance given $w_{t-1}=w$, we have:

$$\begin{aligned}\mathrm{Var}\left[w_t \mid w_{t-1}=w\right] &= E\left[\left(\frac{S_t}{S_{t-1}}w_{t-1}+S_t[\theta+\epsilon_t]-w_{t-1}\right)^2 | w_{t-1}=w\right] \\ &= E\left[\left(S_t[\theta+\epsilon_t-w_{t-1}]\right)^2 \mid w_{t-1}=w\right] \\ &= S_t^2\left[E\left[(w_{t-1}-\theta)^2 \mid w_{t-1}=w\right]+1\right]=S_t^2[S_{t-1}+1] \\ &= S_t^2\frac{S_{t-1}}{S_t}=S_tS_{t-1} \end{aligned} \tag{39}$$

For the unconditional distribution of w_{t-1} we use Eq. (1) and the fact that $m-\mu\sim N(0, 1/\tau_m+1/\tau_\theta)$ to get

$$w_0=\frac{\tau_m m+\tau_\theta\mu}{\tau_m+\tau_\theta}=\mu+\frac{\tau_m}{\tau_m+\tau_\theta}(m-\mu) \tag{40}$$

so $E[w_0]=\mu$ and

$$\mathrm{Var}[w_0]=\left(\frac{\tau_m}{\tau_m+\tau_\theta}\right)^2E\left[(m-\mu)^2\right]=\frac{\tau_m}{\tau_\theta(\tau_m+\tau_\theta)} \tag{41}$$

In general for any $t\in N$

$$E[w_t]=\frac{tE\left[\bar{x}\right]+\tau_mE[m]+\tau_\theta\mu}{t+\tau_m+\tau_\theta}=\frac{t\mu+\tau_m\mu+\tau_\theta\mu}{t+\tau_m+\tau_\theta}=\mu \tag{42}$$

and

$$\begin{aligned}\mathrm{Var}[w_t] &= S_t^2E\left[\left\{t(\bar{x}-\mu)+(\tau_m-\mu)\right\}^2\right] \\ &= S_t^2\left[t^2\left(\frac{1}{t}+\frac{1}{\tau_\theta}\right)+\tau_m^2\left(\frac{1}{\tau_m}+\frac{1}{\tau_\theta}\right)+t\tau_m\frac{1}{\tau_\theta}\right] \\ &= \frac{1}{\tau_m}S_t^2\left[t\tau_\theta+t^2+\tau_m\tau_\theta+\tau_m^2+t\tau_m\right] \\ &= \frac{1}{\tau_m}\left[1-S_t^2(t+\tau_\theta)(\tau_m+\tau_\theta)\right] \end{aligned} \tag{43}$$

which is equivalent to expression (6). □

References

Abowd, J.M., Card, D., 1989. On the covariance structure of earnings and hours changes. Econometrica 57 (2), 447–480, March.

Abowd, J., Kramarz, F., Margolis, D., 1999. High wage workers and high wage firms. Econometrica 67, 251–334.

Abraham, K.G., Farber, H.S., 1987. Job duration, seniority, and earnings. American Economic Review 77 (3), 278–297, June.
Altonji, J.G., Shakotko, R.A., 1987. Do wages rise with job seniority? Review of Economic Studies 54 (3), 437–459, July.
Altonji, J.G., Williams, N., 1997. Do wages rise with job seniority? A reassessment. National Bureau of Economic Research Working Paper 6010, Cambridge, MA (April).
Barron, J.M., Black, D.A., Loewenstein, M.A., 1989. Job-matching and on-the-job training. Journal of Labor Economics 6 (1), 1–19.
Barron, J.M., Berger, M.C., Black, D.A., 1999. Do workers pay for on-the-job training? Journal of Human Resources 34 (2), 235–252.
Becker, G.S., 1975. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. second edn. Columbia University Press, New York.
Ben-Porath, Y., 1967. The production of human capital and the life-cycle of earnings. Journal of Political Economy 75, 352–365.
Chamberlain, G., 1984. Panel data. In: Griliches, Z., Intriligator, M. (Eds.), Handbook of Econometrics, vol. II. North Holland, Amsterdam.
Farber, H.S., 1994. The analysis of interfirm worker mobility. Journal of Labor Economics 12 (4), 554–593, October.
Farber, H.S., 1999. Mobility and stability: the dynamics of job change in labor markets. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3B. North Holland, Amsterdam, Chapter 37.
Farber, H.S., Gibbons, R., 1996. Learning and wage dynamics. Quarterly Journal of Economics, 1007–1048.
Flinn, C., 1986. Wages and job mobility of young workers. Journal of Political Economy 94, S88–S110.
Gallant, R., Jorgenson, D., 1979. Statistical inference for a system of simultaneous, implicit equations in the context of instrumental variable estimation. Journal of Econometrics 11, 275–302.
Hause, J.C., 1980. The fine structure of earnings. Econometrica 48, 1013–1029.
Johnson, N., Kotz, S., 1970. Distributions in Statistics, Continuous Univariate Distributions—2. Wiley, New York.
Jovanovic, B., 1979a. Firm specific capital and turnover. Journal of Political Economy 87, 1246–1260.
Jovanovic, B., 1979b. Job matching and the theory of turnover. Journal of Political Economy 87, 972–990.
Jovanovic, B., 1984. Matching, turnover, and unemployment. Journal of Political Economy 92, 108–122.
Kearl, J., 1988. The covariance structure of earnings and income, compensatory behavior, and on-the-job investments. Review of Economics and Statistics 70, 214–223.
Light, A., McGarry, K., 1998. Job change patterns and the wage patterns of young men. Review of Economics and Statistics 80, 276–286.
Loewenstein, M.A., Spletzer, J.R., 1997. Delayed formal on-the-job training. Industrial and Labor Relations Review 51 (1), 82–99.
MacLeod, W.B., Malcomson, J.M., 1993. Specific investment and wage profiles in labour markets. European Economic Review 37 (2/3), 343–354, April.
Marshall, R.C., Zarkin, G.A., 1987. The effect of job tenure on wage offers. Journal of Labor Economics 5 (3), 301–324.
McCall, J., 1970. Economics of information and job search. Quarterly Journal of Economics 84, 113–126.
Mincer, J., 1974. Schooling, Experience and Earnings. University of Columbia Press for the National Bureau of Economic Research, New York.
Mincer, J., Jovanovic, B., 1981. Labor mobility and wages. In: Rosen, S. (Ed.), Studies in Labor Markets. University of Chicago Press, Chicago, pp. 21–63.
Mortensen, D.T., 1988. Wages, separations, and job tenure: on-the-job specific training or matching? Journal of Labor Economics 6 (4), 445–471.
Neal, D., 1995. Industry-specific human capital: evidence from displaced workers. Journal of Labor Economics 13 (4), 653–677.
Parent, D., 2000. Industry-specific capital and the wage profile: evidence from the national longitudinal survey of youth and the panel study of income dynamics. Journal of Labor Economics 18 (2), 306–323, April.
Rosen, S., 1977. Human capital: a survey of empirical research. In: Ehrenberg, R. (Ed.), Research in Labor Economics, vol. 1. JAI Press, Stamford, CT.

Rothschild, M., 1973. Models of market organization with imperfect information: a survey. American Economic Review 81, 1283–1308.
Stigler, G., 1961. The economics of information. Journal of Political Economy 81, 213–225.
Topel, R., 1991. Specific capital, mobility, and wages: wages rise with job seniority. Journal of Political Economy 99 (1), 145–176, February.
Topel, R., Ward, M., 1992. Job mobility and the careers of young men. Quarterly Journal of Economics 107, 439–479.