

Specific Training Sometimes Cuts Wages and Always Cuts Turnover

Lalith Munasinghe, *Barnard College, Columbia University*

Brendan O’Flaherty, *Columbia University*

Turnover falls with tenure, but wages do not always rise (and sometimes fall) with tenure. We reconcile these findings by revisiting an old issue: how gains from firm-specific training are split between workers and firms. The division is determined by a stationary distribution of outside offers. The lower the wage a firm pays to a specifically trained worker, the more profit it makes but the more likely the employee is to leave. The optimal time paths of wages and turnover show that, if marginal product is increasing, wages need not be increasing but it always implies a falling turnover rate.

I. Introduction

Most theories of the labor market predict that, if specific training occurs, increasing tenure on a job will both raise wages monotonically and cut turnover. Since current empirical evidence supports very strongly the proposition that tenure cuts turnover but does not support the proposition that tenure raises wages, we are left in a quandary: extant theory cannot tell us whether specific training occurs or not. In this article, we resolve this quandary with a new theory. If specific training occurs, increasing tenure must cut turnover but it need not always raise wages. Wages that do not rise monotonically do not indicate that specific training is absent.

The positive cross-sectional association between tenure and earnings orig-

We thank Emanuele Gerratana, Cynthia Howells, Tack-Seung Jun, and Nachum Sicherman for many helpful comments. Contact the corresponding author, Lalith Munasinghe, at lm25@columbia.edu.

[*Journal of Labor Economics*, 2005, vol. 23, no. 2]
© 2005 by The University of Chicago. All rights reserved.
0734-306X/2005/2302-0001\$10.00

inally led researchers to conclude that wages rose with tenure. However, more recent evidence based on longitudinal data shows that the positive wage returns to tenure are small (Abraham and Farber 1987; Altonji and Shakotko 1987; Abowd, Kramarz, and Margolis 1999), despite an ubiquitous strong negative tenure effect on turnover.¹ Within-job wage cuts are not uncommon (Baker, Gibbs, and Holmstrom 1994), and some evidence even points to negative returns to seniority (Ransom 1993). In summary, this empirical literature shows that wage dynamics are more complicated and variegated than the simple presumption that wages monotonically increase with job tenure.

We reconcile these findings by revisiting an old issue: how the gains from specific training are split between workers and firms. The division is determined by a stationary distribution of outside offers. The model is *ex post* monopsony: the lower the wage a firm pays to a specifically trained worker, the more profit it makes and the more eager it is to have that employee stay, but also the more likely the employee is to leave. We solve for the optimal time paths of wages and turnover probabilities. Even if marginal product is increasing, wages need not be increasing; however, rising marginal product always implies a falling turnover rate. While specific training should always cut turnover, there is no theoretical reason why it should always raise wages. We show why sometimes wage cuts happen. Wages eventually rise if you stay employed forever, but nothing in our model guarantees that your completed wage profile will be upward sloping. Whenever there is a wage cut, there is a probability that the worker will leave in that period.

Our model is designed to address some leading empirical puzzles in labor economics. The consensus among empirical labor economists is that wage returns to tenure are small—not that they are zero. So wages on average do increase with tenure, despite occasional within-job wage cuts. The well-recognized puzzle in the literature is why an additional year of tenure reduces turnover by a lot but increases wages by very little. Our results are consistent with this asymmetric tenure effect on wages and turnover. Moreover, turnover decreases are compatible even in the presence of wage cuts. Thus, our model provides an explanation for monotonically decreasing turnover, weak but positive tenure effects on wages, and, especially, occasional wage cuts.²

The remainder of this article is organized as follows. In Section II, we discuss some related theory and highlight various features of the model.

¹ Topel (1991) shows that wages do rise substantially with seniority. However, in a reassessment, Altonji and Williams (1997) argue that Topel overestimates these returns.

² Demand shocks and nominal wage rigidities are other obvious explanations for real wage cuts. Our point is that within-job wage cuts could also occur for less obvious reasons.

Section III presents the model and shows that turnover always falls as tenure increases. Section IV derives our main result: wage cuts can occur even if marginal product is increasing. We show the conditions (on the outside offer distribution) under which wage cuts will occur, and we argue that there is nothing unusual or bizarre about such conditions. The article concludes, in Section V, with a brief summary.

II. Related Theory

The problem of wage determination when an employment relationship generates firm-specific rents has been well known since Becker's original idea of sharing costs and returns of firm-specific investments as a means of providing mutual insurance to each party's investment (Becker 1962; Parsons 1972; Hashimoto 1981). These models show the dual effects of tenure, namely, wage increases and turnover decreases. Our model also predicts turnover decreases, but, unlike these earlier models, it predicts that wages do not increase monotonically with tenure even if marginal product increases monotonically with tenure.

As is well known, Becker-type sharing rules imply inefficiently high turnover because the party initiating turnover does not take into consideration the resulting loss to the other party's investment. Our assumption of ex post monopsony also generates inefficiently high turnover. Mortensen (1978) considers various employment agreements, including "matching alternative offers," that would induce both workers and employers to pursue efficient or joint wealth-maximizing search strategies. Although offer matching has been widely adopted as a wage-setting mechanism since Mortensen (e.g., Harris and Holmstrom 1982; Postel-Vinay and Robin 2002; Munasinghe 2005), we exclude Mortensen-type ex post bargaining in our model. So the question is whether more efficient wage policies such as offer matching can also generate wage cuts. We argue that wage cuts are not an artifact of inefficiency. We show that a model of efficient turnover—offer matching à la Mortensen—leads to both occasional wage cuts and decreasing turnover. However, various considerations make offer matching somewhat controversial.

Jovanovic's article "Firm-Specific Capital and Turnover" (1979), like our model, integrates specific human capital theory and job search theory. Jovanovic shows how match quality jointly determines search effort and investment in specific human capital. Jovanovic's central result is that turnover declines with tenure. He derives this result under the restrictive assumption that wages always equal marginal product. We show the same result under the weaker condition that firms have no commitment to pay marginal product every period. Clearly, the wage sequence a firm can offer a worker depends on the commitment ability of the firm—that is, whether a firm can be trusted not to renege on a promised wage schedule,

since productivity increases on the job are assumed to be firm specific. Because we focus on the no-commitment-ability case—that is, on self-enforcing contracts—our results, unlike Jovanovic's, are time consistent. However, we also show that, if firms have complete commitment ability, then they will delay payments for as long as they can. This is an argument for pensions quite independent of disciplinary considerations.

The model presented here is also closely related to another paper (Munasinghe 2005) that was designed to explain a variety of stylized findings on wage and turnover dynamics. Both models build on a common set of human capital and job search assumptions: specific training is strictly learning-by-doing, and thus neither the firm nor the worker needs to make either decisions or sacrifices in terms of investments;³ workers receive a single job offer from a stationary distribution in each period;⁴ and, because the market for prospective workers is competitive, firms offer zero-profit wage schedules.⁵ But, unlike the model here, Munasinghe assumes downward wage rigidity and ex post bargaining. As a result, his model cannot theoretically address the finding about within-job wage decreases.

Our assumption that specific training is strictly learning-by-doing contrasts with other compensation models, such as up-or-out contracts (Kahn and Huberman 1988) and promotion-based contracts (Prendergast 1993), both of which are designed to create incentives for workers to invest in (nonverifiable) firm-specific human capital. The common implication of these models is that the worker is paid a higher posttraining wage, which implies wage increases with job tenure.⁶ The objective of this literature, however, is to understand labor practices as commitment devices to avoid firm opportunism ex post in capturing rents from specific training. The “no-commitment” wage policies, like the one presented here, are relevant to employment relationships when firms cannot rely only on their reputation to be credible.

III. Turnover

A. The Model

Time is discrete, and t designates the number of periods of completed tenure that a worker has accumulated at the beginning of the current

³ As a consequence, both models are silent about layoffs.

⁴ The finite number of offers per period implies fixed search costs, and hence search effort is not endogenous.

⁵ Munasinghe (2005) presents a more complete discussion of how such competitive market might arise when employment relationships generate specific rents.

⁶ Postinvestment wages may not increase by much for reasons given in Waldman (1990). Waldman argues that up-or-out contracts are incentive compatible in an environment where human capital is general but information is private. The return for being retained is not the result of a precommitment to a high postinvestment wage, but rather it is the result of the positive signal (to other firms) and bidding by other firms.

period. Let y_t denote the marginal revenue product of a worker who has accumulated t periods on the job; thus, y_0 is the marginal product of a new worker. We assume that (y_t) is a monotonically increasing sequence; specific training occurs.

Let w_t denote the wage of a worker who has completed t periods; we refer to this as the wage of an age- t worker. At the beginning of each period (except period 0), the worker receives an outside offer. The outside offer has a present value of θ , where θ is a random variable drawn independently each period from the same distribution, for which $G(\cdot)$ is the cumulative distribution function (CDF). (If the worker would get a signing bonus for taking the outside job, that signing bonus is included in θ , along with the expected present value of wages and transitions opportunities.) We assume that $G(\cdot)$ is strictly increasing on its support. Stationarity reflects the assumption that it is specific training that we are talking about; since the training affects only the worker's productivity within the firm, there is no reason for the distribution of outside offers to change.

Let v_t denote the worker's value of optimal continuation with the firm. This value is evaluated before the outside option becomes known. Assume that the discount factor is $\delta < 1$ and that work involves no utility or disutility. Then,

$$v_t = E[\max(\theta, w_t + \delta v_{t+1})], \quad \forall t \geq 1,$$

and

$$v_0 = w_0 + \delta v_1,$$

because the worker cannot receive an outside offer before going to work (we take this as a definition of what an "outside" offer is). It will be convenient to write

$$x_t = w_t + \delta v_{t+1}.$$

Then, the probability that a worker quits depends on x_t ; specifically, a worker stays in period t if and only if $\theta \leq x_t$ (in the event of ties we assume that the worker stays). We call x_t the worker's "comparison value."

Let h_t , $\forall t \geq 1$, denote the probability that an age- t worker stays on the job. Then,

$$h_t = G(x_t).$$

We refer to h_t as the continuation hazard. Hence,

$$v_t = h_t x_t + (1 - h_t) E(\theta | \theta > x_t).$$

For future reference, note that

$$v_t = \int x_t dG(\theta) + \int_{x_t} \theta dG(\theta),$$

and so

$$dv_t = G(x_t)dx_t > 0. \quad (1)$$

On the firm side, let V_t denote the optimal continuation value for a firm with an age- t worker. Each period the firm chooses a wage offer w_t to maximize its expected present value. Since the firm has no precommitment power, it can assume that its future wages, w_{t+1} , w_{t+2} , and so forth, will also be set optimally and dynamically consistently, and so we can consider them as given when it chooses w_t . Thus, we can think of the firm as choosing x_t rather than w_t . But since $G(\cdot)$ is strictly increasing, it is an invertible function, and we can think of the firm as choosing b , the continuation hazard. Let $F(\cdot)$ be the inverse of G :

$$b = G(F(b)).$$

Then, V_t is given by the fundamental recursion equation

$$V_t = \max_b [b(y_t - w_t + \delta V_{t+1}) + (1 - b)V_0], \quad \forall t \geq 1,$$

where

$$w_t = F(b_t) - \delta v_{t+1}.$$

We can further simplify by writing R_t as the total rent from the relationship,

$$R_t = V_t + v_t.$$

Then, the above recursion equation becomes

$$V_t = \max_b [b(y_t - F(b) + \delta R_{t+1}) + (1 - b)V_0].$$

Assume that $V_0 = 0$. We make this assumption because we want to study ex post monopsony—that is, monopsony after the worker has been hired and has acquired some specific skills—not ex ante monopsony as well. So we must suppose that there are many firms identical to this one bidding for the worker and that, if V_0 were positive, even more firms would enter. The market for prospective workers is competitive. This assumption is not crucial, but it simplifies the mathematics and is consistent with much other theory in labor economics.

Finally, let R_t^* denote expected surplus at time t contingent on the worker staying on the job that period. Hence,

$$R_t^* = y_t + \delta R_{t+1}.$$

Note that R_t^* is independent of the firm's action at t .

Then, the recursion equation for the firm is

$$V_t = \max_b [b(R_t^* - F(b))].$$

Let b_t denote the optimal continuation hazard. Then,

$$\begin{aligned} R_t &= V_t + v_t \\ &= b_t R_t^* + (1 - b_t) E(\theta | \theta > F(b_t)) \\ &= b_t R_t^* + \int \theta dG(\theta). \end{aligned}$$

Bargaining and wage setting are somewhat different in period 0. Since we assume as a convention that no outside offers are possible in period 0 (since there is no "outside" until period 0 has been completed), we take the "continuation hazard" to be undefined, since there is no relationship to be continued. A signing bonus (or penalty) has to be set to lure the worker in, not to keep her from leaving, and our perfect competition assumption implies that the bonus/penalty is high enough that V_0 , the firm's value of a new worker, is zero.⁷ Hence,

$$w_0 = y_0 + \delta V_1.$$

Since the signing bonus/penalty is set by a different process from wages in other periods, we will not pay much attention to it. We do not consider it a wage payment.

⁷ This signing bonus plays a similar role to the front-loaded wage contracts in Black and Loewenstein (1991). Their model is based on heterogeneity of mobility costs, and employers have monopsony power because it is costly for workers to switch employers. Although the source of monopsony power in our model is specific training, in both cases, workers, in anticipation of future monopsony power of the employer, will demand higher wages up front. Another interesting point is Black and Loewenstein's claim that, in the absence of specific training, wages are a decreasing function of tenure. However, in their model, specific training tends to increase wages as tenure lengthens. By contrast, as we will see in the next section, our model can generate falling wages even when specific training is ongoing, depending on the distribution of outside offers. The requirement that the value of an incipient relationship to an employer be zero rules out the sort of equilibrium where workers always go hopping quickly from employer to employer, picking up huge signing bonuses with every hop. Employers would not pay such bonuses.

B. General Results

Now we prove that (b_t) is a weakly increasing sequence: turnover always falls as tenure increases.

PROPOSITION 1. The sequence (b_t) is (weakly) increasing.

Proof of proposition 1. See the appendix.

Thus, in our model, no matter what the distribution of outside offers is, turnover decreases as tenure lengthens. The value of the job to the worker, $F(b_t)$, also monotonically increases as tenure grows.

This does not imply, however, that wages grow monotonically. Current wages are only part of the worker's inducement to reject outside offers. The other part is the continuation value. The continuation value may or may not be growing faster than the total inducement that the firm wants to provide. Only if the total inducement is growing faster than the continuation value will wages rise. This may or may not happen. Section IV analyzes this issue.

We can also prove that turnover will be inefficiently high in this equilibrium. This result, however, is well known, intuitively obvious (since it is a monopsony), and the proof would introduce more notation. We therefore omit it.

Turnover in this model depends on R_t^* , the rent the relationship would generate if the worker turned down this period's outside offer. Specific training decreases turnover because it increases rent. More specific training implies lower turnover in the sense that a job where marginal productivity is higher every period than it is in another job will have (weakly) lower turnover every period. Rent drives turnover in Mortensen's (1978) matching-offers model, too.

The relationship between rent and turnover is first order, however, not second order: bigger increases in R_t^* do not necessarily correspond with bigger increases in b_t . This is because the function that links b_t and R_t^* is, in general, nonlinear. Mortensen's model also has this property.

In one case, however, we can derive a result that resembles "more quickly rising marginal product implies more quickly falling turnover rate." Let (y_t') and (y_t'') be two sequences of marginal products, and let (b_t') and (b_t'') , respectively, be the associated sequences of retention rates, all other parameters being the same. The following corollary follows.

COROLLARY 1. Let $y_t' = y_t''$ for all $t \geq T > 1$, and let $y_t' > y_t''$ for all $t < T$. Then, $b_t' = b_t''$ for all $t \geq T$ and $b_t' \geq b_t''$ for all $t < T$, with strict inequality if $b_t'' < 1$.

Proof of corollary 1. See the appendix.

C. Other Closely Related Models

The sequence of wages that the firm offers depends on its commitment ability. We have focused on the case where the firm has no commitment

ability. However, if the firm has complete commitment ability, then the solution to the firm's problem is trivial: it postpones payment as long as it can. In the appendix, we show this for the uniform case.

Another distinctive feature of our model is the inefficiency of ex post monopsony. One alternative model with efficient turnover is Mortensen's (1978) model of offer matching. It is easy to see that this model will also have declining turnover with tenure. The greater the value of rent, the higher the outside offer the employer will match and so the smaller the probability that the worker will leave.

IV. Wages

A. Motivating Example: Two-Point Distribution

The intuition for why wage cuts sometimes occur is best imparted by considering a particular example from one of the simplest distributions of outside offers. The next section shows how this example generalizes.

Suppose, then, that the outside offer can take one of only two values: θ_l with probability p and θ_h with probability $1 - p$. The sequence of marginal products (y_t) is strictly increasing, but the differences between successive values are small in a sense that we will make more precise shortly.

Assuming that marginal products are high enough that the employer would hire the worker in the first place but not so high that the employer would never take a chance on the worker's departure, it would be easy to see what the sequence of wages would have to accomplish. When the worker is relatively junior and so not especially productive, the employer will be prepared to let her go when she receives the high outside offer θ_h . Thus, with junior workers, the employer will set the worker's comparison value x_t at θ_l —so that the junior worker will stay when she receives the low offer θ_l and leave when she receives the high offer θ_h .

However, when the worker is senior and sufficiently productive, the employer will want her to stay no matter what. So the employer will set $x_t = \theta_h$, and the worker will stay forever.

At some point, the worker will transition from being junior (and dispensable) to senior (and indispensable). We call the transition time T . Specifically, T is the first period for which $x_t = \theta_h$ and we consider the worker to be senior. It is easy to calculate T , but we will not do so here. Our implicit assumption about slowly rising marginal product can now be translated into the explicit assumption that $T > 2$. This informal argument can be summarized in figure 1, which graphs x_t as a function of the worker's tenure on the job.

The employer keeps the worker's comparison value at θ_l up until time T and then increases it to θ_h . (Since the worker's termination hazard is p before T and 0 at T and after, this example clearly shows declining turnover.)

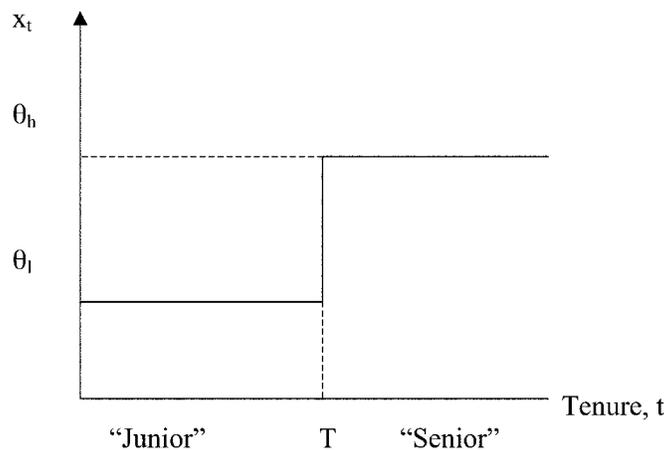


FIG. 1.—Comparison value (x_t) for two-point distribution

What sequence of wages (w_t) implements the sequence of comparison values (x_t) in figure 1? A moment's thought shows that holding wages at a low value before T and then raising them will not create the sequence of comparison values that the employer wants. Such a sequence of wages would make continuation values v_t start rising well before T , and this would mean that comparison values x_t would rise gradually for a junior worker, instead of remaining constant. With constant wages before a jump up, the job gets more valuable as the worker approaches the jump—contrary to the optimal pattern shown in figure 1.

Implementing the sequence of continuation values in figure 1 therefore requires that wages go down from period $T - 2$ to $T - 1$. The wage cut in $T - 1$ must be just big enough that the prospect of losing wages at $T - 1$ and then getting a raise at T and beyond is just the same in present value as getting forever the regular junior wages (the wages at $T - 2$ and earlier). Then as a junior worker approaches T , she becomes no better off—the prospect of the wage rise at $T - 1$ is exactly offset by the prospect of the wage cut at $T - 1$. This creates figure 1's flat profile of comparison values that we know is optimal.

Algebra confirms this verbal argument. For a senior worker, with $t \geq T$, $v_{t+1} = \theta_h$, since continued employment is certain. For a junior worker not approaching the raise, that is, for $t \leq T - 2$,

$$v_{t+1} = p\theta_l + (1 - p)\theta_h \triangleq \bar{\theta},$$

since she will continue if she gets the low outside offer (with probability p) or leave if she gets the high outside offer (with probability $1 - p$). Thus,

$$w_t = x_t - \delta v_{t+1} = \theta_t - \delta \bar{\theta}.$$

But for a junior worker about to get the raise, $t = T - 1$, $v_{t+1} = v_T = \theta_b$, and so

$$w_{T-1} = x_{T-1} - \delta v_T = \theta_t - \delta \theta_b,$$

which is smaller than all preceding wages. To maintain the same termination hazard, the employer has to offer less in current wages to a junior worker about to be promoted.

B. How Special Is This Example?

What was it about the two-point distribution that made wages decrease? In terms of fundamentals, we will argue in this section that a steep CDF of outside offers is the driving force behind this result. The argument in this section will again be graphical and intuitive.

To make this argument, we must take a step back from figure 1 and examine the employer's decision problem in a bit more detail. At any period t , we can consider the problem in (θ, b) space. First, since value is $b(R_{t-1}^* - x)$, once we know R_{t-1}^* , we can draw iso-value curves, as in figure 2. These curves slope upward and asymptote at $x = R_{t-1}^*$. Of course, the employer cannot choose b and x independently: he is bound by the constraint $b = G(x)$, which we illustrate for the two-point case in figure 3. Thus, when R_{t-1}^* is fairly small (but larger than θ_l), the employer's optimum

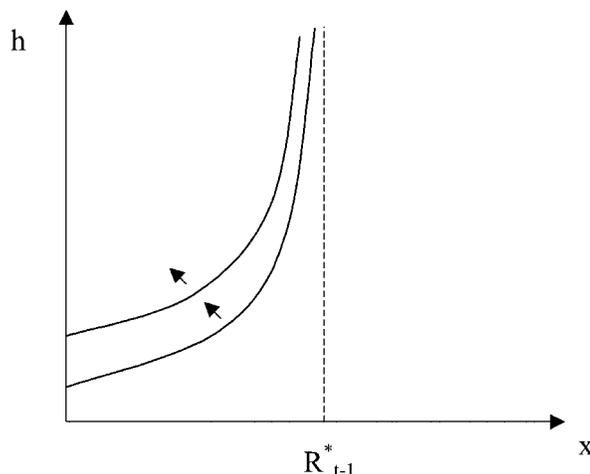


FIG. 2.—Iso-value curves for given R_{t-1}^*

is at the lower kink, as we show in figure 4 (and we consider the worker junior). Small changes in R_{t-1}^* do not change the employer's optimum, and so long as the worker stays junior, x_t does not change. That is the nature of a kink in a constrained optimization program (and the fact that the kink in this example is a right angle is not necessary). If $G(\cdot)$, for instance, looked as in figure 5, the employer would still stay on the lower kink for many periods if marginal product increased slowly.

Now consider how the first difference in wages relates to the first difference in comparison values. From the definition of comparison value,

$$w_{t+1} - w_t = (x_{t+1} - x_t) - \delta(v_{t+2} - v_{t+1}).$$

Notice that the second term in (1) is nonpositive, since worker's value is weakly monotonically increasing. This term, moreover, depends solely on events occurring after $t + 1$. Thus, wages can rise only if the difference $(x_{t+1} - x_t)$ is sufficiently great. But when the kink captures comparison value at the same place for a long period of time, no force is present to raise wages.

Of course, as long as $(v_{t+2} - v_{t+1}) = 0$, no force is present in the junior phase to reduce wages either, and so they stay constant. But when $t = T - 2$,

$$v_{t+2} - v_{t+1} = v_T - v_{T-1} > 0,$$

and so wages go down.

Thus, the feature of the two-point distribution that makes wage cuts possible is the steepness of the CDF in the vicinity of θ_t that makes successive values of x_t the same for at least two periods. But making successive values of x_t the same was not necessary for the argument we used—all that was needed was for them to be sufficiently small. Thus, kinks in the CDF are not required either—all that is needed is that the CDF be sufficiently steep in some region and sufficiently flat nearby that successive values of x_t be sufficiently close.

Wage cuts, therefore, do not require that the distribution of outside offers be atomic. It is also easy to construct examples with multiple wage cuts (if marginal products rise slowly enough, then if the distribution of outside offers is an n -point distribution, a worker will experience $n - 1$ wage cuts if she stays with the same firm throughout her career). We can even construct examples with wage cuts in arbitrarily many successive periods. (Even if wages are cut in arbitrarily many periods, both the workers and the employer must anticipate that, at some point, wages will rise if the employment relationship continues long enough. In this sense, wage cuts are "temporary." But there is never a guarantee that the relationship will last long enough for wages to rise again.) In the next section, we will generalize this intuition and show more rigorously which sorts of outside offer CDFs lead to wage cuts.

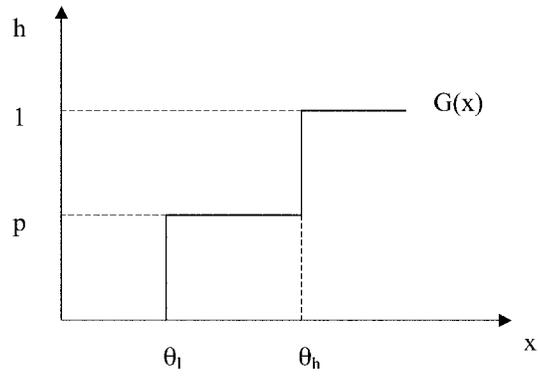


FIG. 3.—Likelihood of staying for two-point distribution

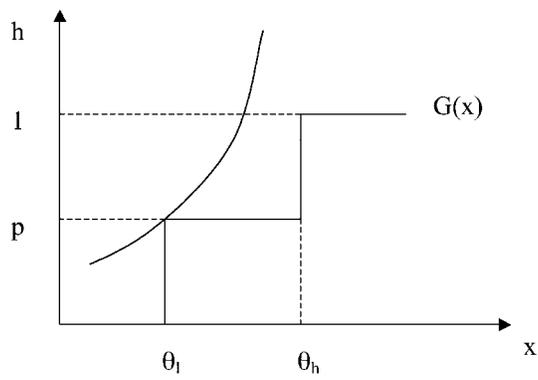


FIG. 4.—Employer's optimum

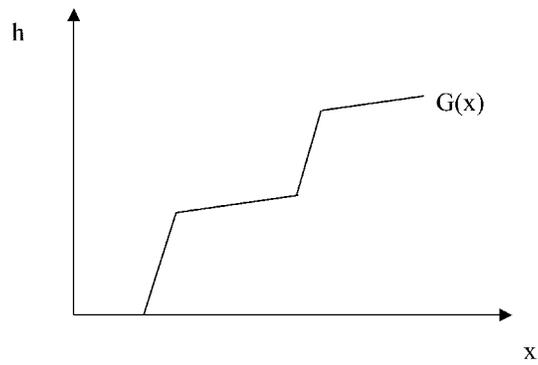


FIG. 5.—Alternative distribution of outside offers

C. General Results on Wage Cuts

Wages will fall when the CDF of the outside offer distribution in the neighborhood of x_t is rising steeply enough, provided that the employment relationship is productive enough and the value of the relationship to the worker will grow quickly enough in the future. In other words, constructing examples with wage cuts is relatively easy and does not require unusual distributions of outside wage offers or bizarre marginal product sequences. Of course, since little or nothing is known empirically about the kinds of outside offer distributions workers actually experience, we cannot make any statements about the probability of wage cuts.

In this section, we will show some sufficient conditions for wages at period $t + 1$ to be lower than wages at period t , for $t \geq 1$. These conditions will not be expressed directly in terms of primitives, but we will show that, for some sets of primitives, they will be satisfied.

In this section, we will assume that the CDF of outside offers $G(\cdot)$ is twice differentiable. We have already shown, with the example of the two-point distribution, that wage cuts are possible when the CDF of outside offers is not differentiable.

Let $g(\cdot)$ denote the probability distribution function (PDF) corresponding to the CDF G . Our main result about wage cuts is given below.

PROPOSITION 2. Let $G(\cdot)$ be twice differentiable for $x < x_{t+1}$ and satisfy

$$[g(x)]^2 > \frac{1}{2}g'(x)G(x), (\forall x). \quad (2)$$

If for all $x \leq x_{t+1}$,

$$\frac{G(x)}{g(x)} < R_t^* - x_{t+1} + \delta(v_{t+2} - v_{t+1}), \quad (3)$$

then $w_t > w_{t+1}$.

Proof of proposition 2. See the appendix.

The intuition behind this proposition, especially the condition shown in (3), is that, if the slope of the CDF G is sufficiently steep in the neighborhood of x_t , as in figure 6, then x_t will be very close to x_{t+1} . (Steepness is reflected in $g(x)$ being large relative to $G(x)$ and so $G(x)/g(x)$ being small.) “Sufficient” steepness in this case depends on how productive the relationship is, R_t^* , and how quickly the value of the relationship to the worker will grow in the future, $\delta(v_{t+2} - v_{t+1})$.

D. Alternative Models

It is trivial to show that, in Mortensen’s offer-matching model, wages will sometime fall. When the worker gets a bad draw from the outside-offer distribution after a good draw, then the employer will cut the value of the job to the worker, not just the wage, even with marginal product

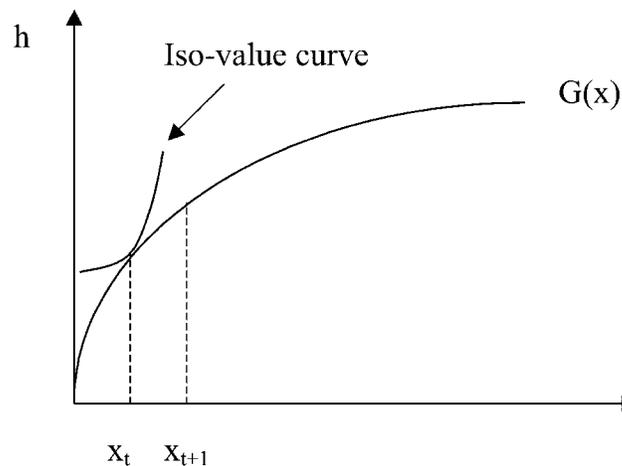


FIG. 6.—Employer's optimum with differentiable cumulative distribution function

increasing. Hence, offer matching, like our model, can generate both wage cuts and always-decreasing turnover.

Of course, wage cuts with offer matching will only occur if outside offers must be accepted or rejected within a finite time period—that is, if offers have time limits.⁸ (If workers report only “good” offers to their current employers, wages will still fall when bad offers are received, since employers can infer from a worker's silence that her offer this period was bad.) If offers last only one period, wages will go down whenever a bad offer follows a good offer. The frequency of wage cuts will, of course, depend inversely on the length of time it takes for an offer to expire. Mortensen implicitly assumes downward wage rigidity despite his claim that outside offers must be accepted or rejected immediately—that is, “no recall” of outside offers.⁹

However, various considerations make offer matching somewhat controversial. First, offer matching (with finite recall) leads to implausible life-cycle wage dynamics. Individual wages over time will follow a mean-reverting process. This runs counter to considerable empirical evidence

⁸ With infinite recall of outside wage offers, offer matching would never generate within-job wage cuts (Munasinghe 2005). But permitting infinite recall does not seem like an always reasonable strategy to follow for employers making outside offers.

⁹ In Mortensen's continuous time framework, current wage remains the same as the previous period wage unless a better outside offer is matched by the firm (or the worker separates).

about how wages evolve within and between jobs over time.¹⁰ By contrast, although our model generates (occasional) wage cuts, job value always increases (because productivity increases), and sooner or later wages must also increase. Second, offer matching may involve high transactions costs (Hashimoto 1981). Third, if information about outside offers is nonverifiable, then offer matching is not feasible, and thus, firms would have to resort to time-consistent wage policies like the one proposed in this article.

If employers have full commitment power, then wage cuts are impossible but the wage-tenure profile is different from empirically observed profiles. The worker is paid only in the last period and never before the last period.

V. Conclusion

We present a model to reconcile recent findings on wage and turnover dynamics. Although the evidence of a positive cross-sectional association between earnings and seniority is widespread, the fairly recent use of longitudinal data has allowed researchers to ask whether individual wages do rise with seniority. Somewhat surprisingly, the evidence on wage changes with tenure appears to be far more variegated than the simple presupposition that wages rise with seniority. Average wage increases with tenure appear to be small, if not negligible, and within-job wage cuts are not uncommon either. The challenge to theory is to reconcile this complex picture of wage dynamics with the fact that the negative relation between turnover and tenure remains as ubiquitous as ever. Our model, built on the workhorse theories of specific training, job search, and self-enforcing contracts, shows that, even if marginal product is increasing due to ongoing specific training, wages need not be increasing; however, rising marginal product always implies a falling turnover rate.

Appendix

Proof of Proposition 1

We proceed by a series of lemmas.

LEMMA 1. V_i is an increasing function of R_i^* .

Proof. Let b_i maximize $b(R_i^* - F(b))$ for $i = 1, 2$ and assume that $R_1^* > R_2^*$. Since b_1 is a maximizer,

$$b_1(R_1^* - F(b_1)) \geq b_2(R_1^* - F(b_2)). \quad (\text{A1})$$

¹⁰ Of course, average wage increases over the life cycle and within jobs could be due to general skill improvements. However, if specific training is important, then the wage dynamic predictions of offer matching with finite recall does not square with the empirical evidence.

Since $b_2 > 0$ and $R_1^* > R_2^*$,

$$b_2(R_1^* - F(b_2)) > b_2(R_2^* - F(b_2)). \quad (\text{A2})$$

And so combining (A1) and (A2),

$$b_1(R_1^* - F(b_1)) > b_2(R_2^* - F(b_2)).$$

LEMMA 2. h_i is a weakly increasing function of R_i^* .

Proof. Retain the same notation as in the proof of lemma 1. Proof is by contradiction. Suppose that $b_2 > b_1$. From (A1),

$$b_2F(b_2) - b_1F(b_1) \geq R_1^*(b_2 - b_1).$$

Then

$$R_1^* \leq \frac{b_2F(b_2) - b_1F(b_1)}{b_2 - b_1}. \quad (\text{A3})$$

Similarly, since b_2 is a maximizer,

$$b_2(R_2^* - F(b_2)) \geq b_1(R_2^* - F(b_1)),$$

and so

$$(b_2 - b_1)R_2^* \geq b_2F(b_2) - b_1F(b_1).$$

Thus,

$$R_2^* \leq \frac{b_2F(b_2) - b_1F(b_1)}{b_2 - b_1}. \quad (\text{A4})$$

Combining (A3) and (A4), we obtain

$$R_2^* \geq R_1^*,$$

which is a contradiction.

LEMMA 3. R_i is a weakly increasing function of R_i^* .

Proof. Retain the same notation as lemmas 1 and 2, and let v_i denote the corresponding continuation value, $i = 1, 2$ and $x_i = F(b_i)$, $i = 1, 2$. Since $R_i = V_i + v_i$ and by lemma 1 we have shown that V_i is a weakly increasing function of R_i^* , we need only know that v_i is a weakly increasing function of R_i^* . Note that lemma 2 implies that $x_1 \geq x_2$. Then,

$$\begin{aligned} v_1 - v_2 &= \int_{x_1}^{x_1} x_1 dG(\theta) + \int_{x_1}^{\theta} \theta dG(\theta) - \int_{x_2}^{x_2} x_2 dG(\theta) - \int_{x_2}^{\theta} \theta dG(\theta) \\ &= \int_{x_2}^{x_1} (x_1 - x_2) dG(\theta) + \int_{x_2}^{x_1} (x_1 - \theta) dG(\theta) + \int_{x_1}^{\theta} (\theta - \theta) dG(\theta) \\ &\geq 0, \end{aligned}$$

since the first two integrals are nonnegative and the third is zero.

LEMMA 4. An increase in $y_{t+\tau}$ for any $\tau \geq 0$ increases R_t .

Proof. Let $\tau = 0$. If y_t increases, R_t^* increases, and by lemma 2, R_t increases.

Let $\tau > 0$. If $y_{t+\tau}$ increases, $R_{t+\tau}^*$ increases. By lemma 3, the increase in $R_{t+\tau}^*$ increases $R_{t+\tau}$. That implies that

$$R_{t+\tau-1}^* = y_{t+\tau-1} + \delta R_{t+\tau}$$

increases. The increase in $R_{t+\tau-1}^*$, in turn, increases $R_{t+\tau-1}$, which increases $R_{t+\tau-2}^*$, and so on, until R_t increases.

Thus we can write

$$R_t = \Omega(y_t, y_{t+1}, y_{t+2}, \dots, y_{t+\tau}, \dots),$$

given G and δ , and note that $\Omega(\cdot)$ is a weakly increasing function of each of its arguments.

LEMMA 5. (R_t) is a (weakly) increasing sequence.

Proof. Consider R_t and R_{t+1} .

$$R_t = \Omega(y_t, y_{t+1}, \dots, y_{t+\tau}, \dots),$$

$$R_{t+1} = \Omega(y_{t+1}, y_{t+2}, \dots, y_{t+\tau+1}, \dots).$$

Since $y_{t+1} > y_t$, $y_{t+2} > y_{t+1}$, $y_{t+\tau+1} > y_{t+\tau}$, \dots , $R_{t+1} \geq R_t$, by repeated application of lemma 4.

Proof of proposition. By assumption, (y_t) is a weakly increasing sequence. By lemma 5, R_t is a weakly increasing sequence. Hence, since

$$R_t^* = y_t + \delta R_{t+1}$$

is a weakly increasing sequence, by lemma 1, b_t is a weakly increasing sequence also.

Proof of corollary 1. For any t , $0 < t < T$, and τ , $\tau \geq T$, we have

$$y_\tau'' - y_t'' > y_\tau' - y_t',$$

the double-prime marginal product is increasing faster than the single-prime marginal product, and

$$(1 - b_t'') - (1 - b_\tau'') \geq (1 - b_t') - (1 - b_\tau'),$$

the double-prime turnover rate is falling faster than the single-prime turnover rate.

Complete Commitment Case

With complete commitment ability, the solution to the firm's problem is trivial: it postpones payment as long as it can. It is easiest to understand this conclusion if we impose the restriction that no worker can work for more than T^* periods—for biological reasons, say. Then, the firm's problem is to choose a sequence of wages (w_t) , $t = 0, \dots, T^*$, that maximizes the expected present value of its profit. Let V_0 denote this objective func-

tion, evaluated, as it must be, just as a worker is hired. We assume that outside offers are uniformly distributed between 0 and 1.

It is easiest and most intuitive to consider the case $T^* = 1$ first. Then,

$$V_0 = \max_{w_0, w_1} \{(\gamma_0 - w_0) + \delta[w_1(\gamma_1 - w_1) + (1 - w_1)V_0]\}, \quad \forall t \geq 1.$$

Clearly, profit is maximized by setting $w_0 = 0$ (optimal w_1 is more difficult to calculate). Increases in w_0 merely reduce first-period profits; they do not increase period 1 retention because they are just irrelevant history when the worker has to make his period 1 decision.

Now consider $T^* = 2$. Clearly $w_0 = 0$. The probability of not quitting at the beginning of period 1 is

$$x_1 = w_1 + \delta v_2,$$

where v_2 is an increasing function of w_2 and can be made as large as desired up to a maximum of one by making w_2 sufficiently large for any value of w_1 . Conditional on not quitting at the beginning of period 1, the value of additional profit is

$$\pi_1(w_1, w_2) = \gamma_1 - w_1 + \delta[w_2(\gamma_2 - w_2) + (1 - w_2)V_0].$$

But π_1 is clearly maximized by setting $w_1 = 0$, since any desired value of x_1 can be obtained by setting $w_1 = 0$ and making w_2 sufficiently large. Optimal wage scheme has $w_1 = 0$, since profits after period zero depend only on x_1 and π_1 .

By repeated arguments like this, we can establish that optimal wage schemes when the firm has unlimited commitment power have positive payments only in the last period; these schemes have maximum retentive power for any expected outlay. Clearly, this result holds for any distribution of outside offers, not just the uniform.

Proof of Proposition 2

Consider the employer's problem at period t :

$$\max_x G(x)[R_t^* - x].$$

The first-order condition for a maximum is

$$g(x)(R_t^* - x) - G(x) = 0,$$

or

$$\frac{G(x)}{g(x)} = R_t^* - x. \quad (\text{A5})$$

The second-order condition is

$$g'(x)(R_t^* - x) - 2g(x) < 0.$$

Substituting from (A5), this becomes

$$g'(x) \frac{G(x)}{g(x)} - 2g(x) < 0,$$

or

$$[g(x)]^2 > \frac{1}{2} g'(x) G(x),$$

which is condition (2) in proposition 2. Hence, the second-order condition holds, and (A5) is necessary and sufficient for a maximum of the employer's problem. Hence,

$$\frac{G(x)}{g(x)} = R_t^* - x_t,$$

or

$$x_t = R_t^* - \frac{G(x)}{g(x)}. \quad (\text{A6})$$

Substituting in (1),

$$\begin{aligned} \omega_{t+1} - \omega_t &= \left(x_{t+1} - R_t^* + \frac{G(x)}{g(x)} \right) - \delta(v_{t+2} - v_{t+1}) \\ &= \frac{G(x)}{g(x)} - \{R_t^* - x_{t+1} + \delta(v_{t+2} - v_{t+1})\}. \end{aligned}$$

If condition (3) in proposition 2 holds, this expression is negative. Hence,

$$\omega_t > \omega_{t+1}.$$

References

- Abowd, Joseph, Francis Kramarz, and David Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67, no. 2:251–333.
- Abraham, Katherine, and Henry Farber. 1987. Job duration, seniority, and earnings. *American Economic Review* 77 (June): 278–97.
- Altonji, Joseph, and Robert Shakotko. 1987. Do wages rise with job seniority? *Review of Economic Studies* 54 (July): 437–59.
- Altonji, Joseph, and Nicolas Williams. 1997. Do wages rise with job seniority? A reassessment. Working Paper no. 6010, National Bureau of Economic Research, Cambridge, MA.
- Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994. The wage policy of a firm. *Quarterly Journal of Economics* 109 (November): 921–56.
- Becker, Gary. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70 (October): 9–49.

- Black, Dan, and Mark Loewenstein. 1991. Self-enforcing labor contracts with costly mobility: The subgame perfect solution to the chairman's problem. *Research in Labor Economics* 12:63–83.
- Harris, Milton, and Holmstrom, Bengt. 1982. A theory of wage dynamics. *Review of Economic Studies* 49 (July): 315–33.
- Hashimoto, Masanori. 1981. Firm-specific human capital as a shared investment. *American Economic Review* 71, no. 3:475–82.
- Jovanovic, Boyan. 1979. Firm-specific capital and turnover. *Journal of Political Economy* 87 (December): 1246–60.
- Kahn, Charles, and Gur Huberman. 1988. Two-sided uncertainty and “up-or-out” contracts. *Journal of Labor Economics* 6 (October): 423–44.
- Mortensen, Dale. 1978. Specific capital and labor turnover. *Bell Journal of Economics* 9 (Autumn): 572–86.
- Munasinghe, Lalith. 2005. A theory of wage and turnover dynamics. Unpublished manuscript, Department of Economics, Barnard College, Columbia University.
- Parsons, Donald. 1972. Specific human capital: An application to quit rates and layoff rates. *Journal of Political Economy* 80 (November/December): 1120–43.
- Postel-Vinay, Fabien, and Jean-Marc Robin. 2002. The distribution of earnings in an equilibrium search model with state-dependent offers and counteroffer. *International Economic Review* 43, no. 4:989–1016.
- Prendergast, Canice. 1993. The role of promotion in inducing specific human capital acquisition. *Quarterly Journal of Economics* 108 (May): 523–34.
- Ransom, Michael. 1993. Seniority and monopsony in the academic labor market. *American Economic Review* 83 (March): 221–33.
- Topel, Robert. 1991. Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy* 99, no. 1:145–76.
- Waldman, Michael. 1990. Up-or-out contracts: A signaling perspective. *Journal of Labor Economics* 8 (April): 230–50.