



MIT Sloan School of Management

Working Paper 4324-03

Institute for Work and Employment Research (IWER)

Working Paper 07-2003

June 2003

ENRICHING A THEORY OF WAGE AND PROMOTION DYNAMICS INSIDE FIRMS

Robert Gibbons and Michael Waldman

© 2003 by Robert Gibbons and Michael Waldman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission, provided that full credit including © notice is given to the source.

This paper also can be downloaded without charge from the
Social Science Research Network Electronic Paper Collection:

<http://ssrn.com/abstract=423260>

ENRICHING A THEORY OF WAGE AND
PROMOTION DYNAMICS INSIDE FIRMS

by

Robert Gibbons
MIT and NBER
Cambridge, MA 02142
(617) 253-0283, rgibbons@mit.edu

and

Michael Waldman
Johnson Graduate School of Management
Cornell University
Sage Hall
Ithaca, NY 14853
(607) 255-8631, mw46@cornell.edu

June 2003

*We thank Orley Ashenfelter, George Baker, Gary Fields, Paul Oyer, Jan Zabochnik, seminar participants at Cornell University, Queen's University, UCLA, and Washington University in St. Louis, and conference participants at the 2002 SOLE conference and the 2002 SITE and NBER conferences on personnel economics for helpful comments, and Ari Gerstle and Kameshwari Shankar for research assistance.

ABSTRACT

In previous work we showed that a model that integrates job assignment, human-capital acquisition, and learning can explain several empirical findings concerning wage and promotion dynamics inside firms. In this paper we extend that model in two ways. First, we incorporate schooling into the model and derive a number of testable implications that we then compare with the available empirical evidence. Second, and more important, we show that introducing “task-specific” human capital allows us to produce cohort effects (i.e., the finding that a cohort that enters a firm at a low wage will continue to earn below-average wages years later). We argue that task-specific human capital is a realistic concept and may have many important implications. We also discuss limitations of our (extended) approach.

I. INTRODUCTION

A standard theme in labor economics, industrial relations, and human resource management has been that careers in organizations are inconsistent with standard economic theories of labor markets. For example, Doeringer and Piore (1971) described aspects of internal labor markets (such as ports of entry and wages attached to jobs) that depart from standard conceptions of competitive labor markets. More recently, Medoff and Abraham (1980,1981), Lazear (1992), and Baker, Gibbs, and Holmstrom (1994a, 1994b) provided detailed empirical analyses of careers inside specific firms and argued that many of their results again do not match standard conceptions of competitive labor markets. In particular, Baker, Gibbs, and Holmstrom (1994b) compared their empirical findings to two theories of wage determination - learning and incentives - and found their results to be quite different than the predictions of those theories.

In Gibbons and Waldman (1999a), we developed a theoretical model that integrates job assignment, human-capital acquisition, and learning. Regarding job assignment, we followed Rosen (1982) and Waldman (1984a) by assuming that jobs are ranked in terms of the importance of ability, so it is efficient to assign higher-ability workers to higher levels of the job ladder. Regarding human-capital acquisition, we followed the spirit of Becker (1962) and Mincer (1974) by assuming that a worker's productivity rises during the worker's career. (We also assumed productivity rises faster for workers of higher innate ability.) Finally, regarding learning, we followed Harris and Holmstrom (1982) and Farber and Gibbons (1996) by assuming there is uncertainty about a worker's ability when she enters the labor force, but all firms learn about the worker's ability from public output observations as the worker ages.

Our 1999a model captures many of the empirical results concerning wage and promotion dynamics found by Medoff and Abraham, by Lazear, by Baker, Gibbs, and Holmstrom, and by others.¹ For example, Baker, Gibbs, and Holmstrom (henceforth, BGH) find that (1) wage changes are serially correlated, (2) promotions are associated with large wage increases, and (3) workers who receive large wage increases early in their stay at one level of the job ladder are promoted quickly to the next level. Our model captures the first and third of these findings because of the skill-acquisition process that we assume: a worker who receives a large wage increase today is likely a worker with high ability, so the

¹ Some of the other studies whose empirical findings are consistent with one or more of our theoretical results include Rosenbaum (1984), Murphy (1985), Main, O'Reilly, and Wade (1993), McCue (1996), Baker (1997), and Podolny and Baron (1997). See Gibbons (1997) for a review of this literature through 1997. Also, see Lima and Pereira (2001) for a recent study testing our theory that uses a longitudinal matched employer-employee panel of 74 large manufacturing firms in Portugal. Their results are in general supportive of our framework.

worker's productivity will likely grow quickly in the future, producing both a large wage increase and an early promotion. Our model captures the second finding because of the learning process that we assume: a promoted worker is likely a worker whose most recent output produced positive updating of the beliefs concerning the worker's ability, so the worker's wage will likely increase significantly.² Our 1999a model also explains Medoff and Abraham's findings concerning performance evaluations, provided that supervisors evaluate individuals relative to others with the same labor-market experience.

In this paper we extend our 1999a model in two directions. First, we address issues that we did not consider in our earlier work. In particular, we incorporate schooling into the model and derive a number of testable implications that we then compare with the available empirical evidence. Second, and more important, we show that an important empirical finding that is inconsistent with our 1999a model becomes consistent with our approach if we make a realistic new assumption. In particular, we incorporate task-specific human capital into the model and thereby capture the BGH finding of cohort effects: a cohort that enters a firm at a low wage will earn below-average wages years later.

The first part of the paper incorporates schooling into our previous model in two ways. First, schooling increases a worker's starting level of general-purpose human capital. Second, schooling is correlated with the speed with which the worker accumulates human capital once she is on the job. We show that this new model makes a number of predictions concerning how schooling affects wage and promotion dynamics. To begin, this extended model of course predicts a positive correlation between schooling and the starting wage, since more schooling means a higher starting level of general human capital. But the model also makes other predictions, such as that schooling is positively related to promotion prospects (because a worker with more schooling accumulates human capital more quickly once she is on the job) and that schooling is positively related to the wage even after holding job level and experience fixed (because a worker with more schooling has a higher initial level of general human capital). We compare these predictions with the available evidence and find that, although the evidence is not clear concerning all the predictions, in general the predictions coincide well with the available data.

The second part of the paper considers cohort effects. BGH find that the average wage at which a cohort is hired is positively correlated with the cohort's average wage many years later, even after

² Our 1999a model also explains two other findings in the BGH study. The first is that real-wage decreases are not rare but demotions are. The second is that, even though promotions are associated with large wage increases, wage increases at promotion are small relative to the difference in average wages across levels of the job ladder.

controlling for cohort differences in age, gender, race, and education. In our earlier paper we argued that our framework is inconsistent with this finding, because we believed the finding to be at odds with spot-market contracting, which is a basic feature of our approach. Here we show that we were incorrect in that earlier discussion: introducing task-specific human capital into our framework produces cohort effects even in the presence of spot-market contracting. By task-specific human capital we mean that some of the human capital an individual acquires on the job is specific to the particular tasks being performed, as opposed to being specific to the firm. Hence, when a worker is promoted and given a new set of tasks to perform, some of her acquired human capital goes unutilized on the new job. We show that, if task-specific human capital is an important part of the skill-acquisition process, then cohort effects can arise.

Our view is that the first of these extensions, concerning schooling, is quite straightforward, but nonetheless advances the project we advocated in our *Handbook* chapter, Gibbons and Waldman (1999b) - namely, that theoretical models of wage determination should begin to address broad patterns of evidence rather than isolated stylized facts. By incorporating schooling into our model, we are adding several new findings to the (already fairly large) set of findings that we originally set out to explain.

In contrast to the first extension, we believe the second extension, concerning task-specific human capital, is more novel and potentially quite rich in its theoretical applications and empirical implications. Since Becker's (1962,1964) seminal work, the literature has focused almost exclusively on general-purpose and firm-specific human capital. We believe that task-specific human capital is potentially both as commonplace and as important as these better known types. Clearly, when a worker performs a particular task, it is likely that some of the human capital she acquires is useful for performing that task but of only limited use in the performance of other tasks; that is, there may be task-specific learning by doing.³ By introducing task-specific human capital into our model in order to analyze cohort effects, we hope to set the stage for a variety of future theoretical and empirical exercises concerning issues such as job design, job assignment, labor mobility, labor demand, and even business strategy.⁴

³ See Heckman, Lochner, and Cossa (2002) for a recent empirical study that finds evidence consistent with much of the human-capital accumulation on jobs being due to learning-by-doing.

⁴ As far as we know, Kwan (1998) was the first to analyze what we are calling task-specific human capital. He considers this type of human capital in a model similar to Prendergast's (1993), where wage increases at promotion are used to provide incentives for investment in human capital. Kwan's analysis does not consider our main focus here - namely, the implications of task-specific human capital when cohorts vary in the mix of jobs at which they enter the firm. See also Chaudhury (2002) for a recent analysis that incorporates task-specific human capital, but again the focus is not on cohort effects.

In sum, this paper reinforces the contention in our 1999a paper that a framework that integrates job assignment, human-capital acquisition, and learning can explain a large number of empirical findings concerning wage and promotion dynamics inside firms.⁵ Our interpretation is *not* that these three model elements currently (or could ever) constitute a complete theory of wage and promotion dynamics inside firms. Clearly other elements - ranging from incentives to social norms - are important parts of such a theory. Rather, our interpretation is that, for a wide variety of firms, a complete theory of wage and promotion dynamics is very likely to include our three model elements as integral parts. Notice that we say a “wide variety” of firms as opposed to “all” firms. This is because we believe our theory does not apply well to the type of internal labor market initially described by Doeringer and Piore (1971), involving ports of entry, wages attached to jobs, and little turnover at higher levels of the job ladder. We return to this issue in Section IV, where we discuss limitations of our approach.

II. THE ROLE OF SCHOOLING IN WAGE AND PROMOTION DYNAMICS

In this section we consider the role of schooling in wage and promotion dynamics inside firms. We begin by constructing a model that enriches our 1999a model by incorporating schooling. We then analyze the model and show that it produces several natural predictions concerning the role of schooling in wage and promotion dynamics inside firms. We end the section by discussing the empirical evidence on this topic, which is broadly consistent with the model’s predictions. Our main point in this section is not that this modest enrichment is important in its own right, but rather that enriching the model in this way adds several new findings to the broad pattern of evidence that our approach can explain.

A. The Model

All firms are identical and the only input is labor. A worker’s career lasts T periods, $T \geq 3$. Worker i enters the labor market with a schooling level, denoted S_i , which can take on any integer value between 1 and N . We assume that there is a positive number of workers at each value of S . One simple

⁵ Other papers that integrate these three model elements include Bernhardt (1995) and Jovanovic and Nyarko (1997). In contrast to our approach, Bernhardt’s analysis focuses on asymmetric learning, as in Waldman (1984b), Greenwald (1986), and Gibbons and Katz (1991), where a worker’s current employer receives better information concerning the worker’s ability than prospective employers receive. Jovanovic and Nyarko do not consider the ability of their model to capture a broad set of empirical findings. Bernhardt does, but his model cannot explain a number of important findings in the literature; see Section V of Gibbons and Waldman (1999a) for a detailed discussion of Bernhardt’s analysis.

interpretation is that a worker's schooling level is the number of years the worker spent in school prior to entering the labor market.⁶

Greater schooling makes a worker more productive, but so does greater experience. This distinction between schooling and experience is standard in the literature, but we consider the microfoundations of the effects of schooling and experience in more than the usual level of detail, as follows. Let h_{it} denote the worker's "on-the-job human capital," where

$$(1) \quad \dot{h}_{it} = \lambda_i f(x_{it}).$$

In this specification of on-the-job human capital, λ_i is the worker's ability to learn on the job and x_{it} is the worker's labor-market experience prior to period t (i.e., for a worker in her first period in the labor market, prior experience x_{it} equals zero). In this specification, $f(x_{it})$ is the speed with which on-the-job human capital grows in period t , where $f(0) > 0$, $f' > 0$, and $f'' < 0$.⁷ Thus, a worker with more ability to learn on the job (i.e., higher λ_i) accumulates on-the-job human capital more quickly. We assume that a worker with schooling level S has ability to learn on the job equal to $\lambda_H + B(S)$ with probability p and ability to learn on the job equal to $\lambda_L + B(S)$ with probability $(1-p)$, where $\lambda_H > \lambda_L$ and $B > 0$. That is, schooling is positively correlated with a worker's ability to learn on the job (either because schooling is useful for learning on the job or because there is a positive relationship between schooling and innate ability to learn on the job).

A firm consists of two different jobs, denoted 1 and 2. In our 1999a paper, we allowed for three job levels, but little of the evidence discussed below relies on three job levels, so here we assume only two levels to simplify the analysis. If worker i is assigned to job j in period t then the worker produces

$$(2) \quad y_{ijt} = d_j + G(S_i) + c_j(\lambda_{it} + \eta_{ijt}),$$

where d_j and c_j are constants known to all labor-market participants, $G > 0$ and $G' < 0$, and η_{ijt} is a noise term drawn from a normal distribution with mean 0 and variance σ^2 .⁸ Let h_{it} denote the amount of on-

⁶ With this interpretation, it would probably be more accurate to assume that worker i 's career lasts T_i periods, where $T_i = A - S_i$. But since nothing of significance would be changed by assuming this, for simplicity of exposition we assume that the career lasts T periods for all workers.

⁷ We assume $f(0) > 0$ rather than $f(0) = 0$. One interpretation is that this model is a simplified version of a continuous-time model where production in the first period represents production in the early part of the worker's career. Since, on average, during this early part of the career the worker has a positive amount of labor-market experience, it is natural to assume $f(0) > 0$ rather than $f(0) = 0$.

⁸ We could have assumed that the extra productivity due to a worker's starting level of general-purpose human capital depends on the job the worker is assigned to. That is, rather than this productivity being equal to $G(S_i)$, we

the-job human capital at which a worker is equally productive at jobs 1 and 2. That is, θ_i solves $d_1 + c_1 \theta_i = d_2 + c_2 \theta_i$. We assume $c_2 > c_1 > 0$ and $d_2 < d_1 < 0$, so the efficient assignment for a worker with $\theta_i > \theta^*$ is job 2, while the efficient assignment for a worker with $\theta_i < \theta^*$ is job 1.

We assume that each worker's schooling level is known to all labor-market participants when the worker enters the labor market. In contrast, each worker's value for θ_i is not known by either the firms or the worker, although the probability p and ability levels $\theta_L + B(S)$ and $\theta_H + B(S)$ are common knowledge. Learning about θ_i takes place at the end of each period, when the realization of the worker's output for that period is publicly observed. The presence of the noise term ϵ_{ijt} in (2) implies that learning occurs gradually, as follows.

Let $z_{it} = [(y_{ijt} - d_j - G(S_i)) / c_j] - B(S_i) f(x_{it}) - \theta_i - B(S_i) f(x_{it}) + \epsilon_{ijt}$. That is, z_{it} is the signal about the worker's on-the-job human capital after subtracting off the human capital due to the schooling term $B(S_i)$ that the market extracts from observing the worker's output in period t . We refer to z_{it} as the worker's normalized output from period t (i.e., normalized to abstract from job assignment and schooling) and to $(z_{it-x}, \dots, z_{it-1})$ as the worker's normalized output history at date t . Because the signal z_{it} is independent of job assignment and schooling, there is no difference in the rate of learning across jobs and schooling levels. Let θ_{it}^e denote worker i 's expected ability to learn on the job at date t : $\theta_{it}^e = E(\theta_i | S_i, z_{it-x}, \dots, z_{it-1})$. From θ_{it}^e we can compute the expected on-the-job human capital of worker i in period t :

$$(3) \quad \theta_{it}^e = \theta_i^e f(x_{it}).$$

For example, if worker i enters the labor market in period t with schooling S_i , then $\theta_{it}^e = [p \theta_H + (1-p) \theta_L + B(S_i)] f(0)$. (Recall that x_{it} measures prior labor-market experience, so $x_{it} = 0$ in the worker's first period in the labor market.)

Workers and firms are risk neutral and have a discount rate of zero. There is no cost to workers from changing firms, or to firms from hiring or firing workers. Under these assumptions, there are no benefits to long-term contracts, so we assume that wages are determined by spot-market contracting. Finally, to ease the comparison of the model with the empirical evidence, we restrict attention to wages that are paid in advance of production, as opposed to one-period piece-rate contracts.

could have assumed it equals $G_j(S_i)$ for a worker assigned to job j . Given the additional assumptions $G_2(S_i) \geq G_1(S_i)$ for all S_i and $G_2(S_i) \geq G_1(S_i)$ for all S_i , all the qualitative results that follow would still hold.

There is free entry into production. At the beginning of each period, all firms simultaneously offer each worker a wage for that period. The worker then works for the firm that offers the highest wage. If there are multiple firms tied at the highest wage, the worker chooses randomly among these firms (unless one of these tied firms was the worker's employer in the previous period, in which case the worker remains with that firm). This tie-breaking rule is equivalent to assuming a moving cost that is infinitesimally small; as a result, there is no turnover in equilibrium.

To reduce the number of cases that need to be considered, we restrict the analysis to parameterizations that satisfy the following conditions. First, $[p_H + (1-p_L) + B(N)]f(0) < \dots$, so that it is efficient for each worker to be assigned to job 1 in the first period of the worker's career. Second, $[p_H + B(S)]f(T-1) > [p_L + B(S)]f(T-1)$ for all S , so that for each schooling level it is efficient for workers with the highest ability to learn on the job to be on job 2 by the end of their careers, but for workers with the lowest ability it is efficient to remain on job 1.

As a final point, there are a number of ways this model could be made more realistic. For example, we could start with workers having some knowledge of their own innate abilities and incorporate a cost of schooling, and then have the schooling distribution be derived endogenously, as in Spence (1973). Or, we could assume a worker's schooling level is positively related to the extra productivity due to a worker's starting level of general-purpose human capital $G(S_i)$, but not in a deterministic way. We believe our main results are robust to both changes. The reason we consider the particular specification above is that it allows us to present our main results in a tractable analysis.

B) Analysis

We begin with a general result concerning how workers are assigned to jobs and how they are paid. We then turn our attention to how schooling affects wage and promotion dynamics. The general result is that, as in Gibbons and Waldman (1999a), in each period each worker is assigned to the job that maximizes her expected output in that period and paid a wage equal to that expected output.⁹ Let w_{it} denote the equilibrium wage paid to worker i in period t .

⁹ Throughout the analysis we assume that if a firm is indifferent between assigning a worker to either of the two jobs then the firm assigns the worker to the high-level job. This assumption is not crucial for the analysis but simplifies the statements of some results.

Proposition 1: In equilibrium, job assignments and wages are given by i) and ii).

- i) If $\epsilon_{it} < \epsilon_c$, then worker i is assigned to job 1 in period t and paid $w_{it} = d_1 + G(S_i) + c_1 \epsilon_{it}$.
- ii) If $\epsilon_{it} \geq \epsilon_c$, then worker i is assigned to job 2 in period t and paid $w_{it} = d_2 + G(S_i) + c_2 \epsilon_{it}$.

Proposition 1 tells us that the assignment of workers to jobs follows a simple rule: a worker is assigned to the high-level job if and only if her expected on-the-job human capital exceeds the critical value ϵ_c . Only the expected on-the-job human capital matters, rather than other moments of the distribution, because output in each job is a linear function of on-the-job human capital and the rate of learning is independent of job assignment.

This model is a simple extension of our 1999a model, so it of course reproduces most of the BGH findings captured by our original model. Three such findings follow. First, this model captures the finding of serially correlated wage increases. The reason is that a large wage increase today likely means the worker has high ability to learn on the job, which is positively correlated with a large wage increase tomorrow. Second, this model captures the finding that a large wage increase early in a worker's stay at one level of the job ladder is positively correlated with quick promotion to the next level. The logic here is that, similar to the argument above, a large wage increase early in a worker's stay at job 1 likely means the worker has high ability to learn on the job, which implies that on-the-job human capital will likely grow quickly and induce promotion. Third, this model also captures the finding of large wage increases at promotion. As in our 1999a model, this result follows from a selection effect: a promoted worker will likely be a worker for whom learning resulted in a large increase in expected on-the-job human capital.¹⁰

This model also explains a BGH finding that our previous model did not capture. BGH found that adjacent levels of the job ladder had overlapping wage distributions: the highest paid workers at level L made more than the lowest paid workers at level $L+1$. In our earlier model there were no schooling levels and all workers were ex ante identical, so a worker's wage and job assignment depended solely on ϵ_{it} , so the highest paid worker at level L was paid strictly less than the lowest paid worker at level $L+1$. In this section's model, however, a worker's wage depends on both ϵ_{it} and the worker's schooling level. Consequently, in contrast to our earlier analysis, this model exhibits overlapping wage distributions: the

¹⁰ All of these results follow easily from results in our earlier paper given that in each case we are referring to workers of a given schooling level. We believe that each of the results also holds without controlling for workers' schooling levels, but we have not formally shown this.

highest-paid workers assigned to job 1 are paid more than the lowest-paid workers assigned to job 2. Furthermore, our theory not only explains overlapping wage distributions but does so in a testable way: if we consider pairs of workers on adjacent levels, where the worker assigned to the lower level is paid more than the worker assigned to the higher one, the worker on the lower level should typically have more schooling.¹¹

We now turn our attention to predictions concerning the relationship between schooling and wage and promotion dynamics. Corollary 1 states our (unsurprising) first prediction: schooling is positively related to the starting wage.

Corollary 1: The wage a worker receives in her first period in the labor market is a strictly increasing function of the worker's schooling level.

In this model, a worker's starting wage is an increasing function of her schooling level for two reasons. The first is that more schooling means that a worker enters the labor force with more general-purpose human capital and so is paid more. The second is that, because $B > 0$, a worker with more schooling has a higher expected starting level of on-the-job human capital, $E(\varphi(0))$ (see footnote 7).

The second prediction is that schooling is positively related to job level. Corollary 2 develops this prediction. Let (X,S) denote the proportion of workers of experience X and schooling level S who are assigned to the high-level job.

Corollary 2: Given any (X,S) pair, where $1 \leq X \leq T-1$ and $S \leq N-1$, $(X,S+1) > (X,S)$. Also, the inequality is strict if X is sufficiently close to $T-1$.

¹¹ In Gibbons and Waldman (1999a) we discussed a different extension of our original analysis that captures overlapping wage distributions (see footnote 13 of that paper). In particular, we argued that overlapping wage distributions can be explained by each firm having multiple job ladders, where promotions only occur within ladders and not across. Our argument was that, if the wage distributions for one ladder are shifted relative to another, then an analysis of the aggregate job ladder would yield overlapping wage distributions. We believe the schooling argument above is a superior argument. The reason is that, although we are not familiar with any evidence on the subject, we suspect that when a firm has multiple job ladders there are overlapping wage distributions for each job ladder that makes up the aggregate job ladder, and the schooling argument presented here is consistent with such a finding.

The logic for this result follows from the way that schooling is correlated with the speed of human-capital acquisition. Holding experience constant, workers with more schooling will likely have higher ability to learn on the job. This means that if we consider any experience, X , and any critical value, θ^+ , workers with schooling level $S+1$ will have a higher proportion whose expected on-the-job human capital exceeds θ^+ than will workers with schooling level S . Since any worker whose on-the-job human capital exceeds θ^+ will be assigned to the high-level job, we now have that at any experience level the proportion assigned to the high-level job will be increasing in the schooling level.

Notice that although the corollary is directly a statement about job assignments, it easily translates into a statement about promotions. That is, since all workers in our model start their careers in the low-level job, the corollary tells us that the promotion rate is higher for workers with higher schooling levels, i.e., $(T-1,S)/(T-1)$ is increasing in S . Additionally, although we do not show it formally, a refinement of this prediction is that in empirical analyses of promotion rates the relationship between schooling and promotions will be weaker when other measures of worker ability to learn on the job are included as explanatory variables. The logic here is that in our model schooling affects the probability of promotion through its correlation with the worker's ability to learn on the job. Hence, if one included a noisy measure of this ability in a regression analysis of promotion rates, the prediction is that the measured impact of schooling on promotions should fall.

The third prediction is that wage increases are positively related to schooling. Corollary 3 develops this prediction. Let $W(X,S)$ denote the average wage paid to workers of experience X who have schooling level S .

Corollary 3: For any parameterization, if demotions are sufficiently rare, then given any X , $0 < X < T-2$, $W(X+1,S)-W(X,S)$ is a strictly increasing function of S .

The logic for this result is straightforward if we restrict attention to values of X such that all workers are assigned to the low-level job at experience levels X and $X+1$, because the argument is then similar to why workers with more schooling are typically promoted more quickly. As stated above, at any given experience level, workers with more schooling will on average have a higher ability to learn on the job. This means that, as the workers gain experience, those with more schooling will on average experience

higher growth in on-the-job human capital. In turn, since a worker assigned to the low-level job in adjacent periods will experience a wage change equal to c_1 multiplied by the change in the worker's on-the-job human capital, we have that those with more schooling will on average experience higher wage growth.

The result still holds but the argument becomes more complicated when it is possible that a worker will be on different jobs at experience levels X and $X+1$ as long as demotions are rare (which is typically the case).¹² For example, suppose that all workers are assigned to the low-level job when experience equals X but some workers are assigned to the high-level job when experience equals $X+1$. Now there are two reasons why schooling is positively related to wage increases. The first is that those with more schooling typically experience higher growth in on-the-job human capital. The second concerns the fact that, in contrast to what was true before, a worker's wage change does not always equal c_1 multiplied by the change in on-the-job human capital. Rather, for those promoted, the wage change equals a convex combination of c_1 and c_2 multiplied by the change in on-the-job human capital, where $c_2 > c_1$. This results in a second reason that schooling leads to larger wage increases, because from Corollary 2 we know that higher schooling means that more workers are promoted.

Our last prediction is that schooling is positively related to wages even after controlling for experience and job assignment. Let $W_j(X,S)$ be the average wage paid to workers of experience level X and schooling level S who are assigned to job j . Also, to generate this prediction we now assume that the extra productivity due directly to schooling is given by $G(S) = g(S)$, where $g'(S) > 0$, $g''(S) < 0$, and $g(0) > 0$.

Corollary 4: Holding all other parameters fixed, if β is sufficiently large, then $W_j(X,S+1) - W_j(X,S) > 0$ for all (X,S,j) triplets, where $0 \leq X \leq T-1$, $0 \leq S \leq N-1$, and $j=H,L$.

This result follows because of the effect that a worker's starting level of general-purpose human capital has on productivity. Since this effect does not vary across jobs, it does not affect job assignment, but it does affect productivity and thus wages. In turn, since higher schooling means higher productivity due to

¹² The role of demotions being rare is as follows. When a worker is demoted, instead of the worker's wage decrease being c_2 multiplied by the decrease in the the worker's expected on-the-job human capital, the wage decrease is a convex combination of c_1 and c_2 multiplied by the decrease in the worker's expected on-the-job human capital, i.e., the wage decrease is smaller. The result is that, if there is a much higher frequency of demotions at a lower schooling level, then it is possible that the lower schooling level will have a higher return to experience.

the higher starting level of general-purpose human capital, the presence of this factor serves to make the wage a positive function of schooling even after controlling for experience and job assignment.¹³ Note that there is a complicating factor because a worker's schooling level affects the underlying distribution of true ability to learn on the job, so extra schooling will affect (and possibly reduce) the expected ability to learn on the job of workers with a given experience and job assignment. The restriction that β is large means that the first factor dominates, and therefore that the wage increases with schooling even after controlling for experience and job assignment.

C) Empirical Evidence

We now consider the extent to which the available empirical evidence concerning schooling and wage growth supports our theoretical predictions. Unfortunately, most of the vast literature concerning schooling and wage growth does not speak to any of our predictions. That is, because the goal of previous studies was not to test the theory developed above, much of the existing evidence cannot be used either to refute or to support our predictions. Nevertheless, there are a few studies that have reported results that do address our predictions and, although the evidence is a bit mixed, we feel that overall the results support our theoretical predictions.

We begin with BGH. In our previous paper we discussed a variety of findings from the BGH study, but there were two findings concerning schooling that we did not address. First, BGH find that the average schooling of workers promoted into a job level increases with the level the worker is promoted into. Although this is not exactly what Corollary 2 states, it would be easy to extend the model to additional job levels to capture exactly this prediction. Second, BGH run a wage regression similar to the standard Mincerian human-capital earnings function discussed below but also include the worker's job level. Consistent with our model (Corollary 4), they find that higher values for schooling are associated with higher wages, even after controlling for experience and job level.

¹³ Although the argument is a bit more complicated, the result also holds if, as we briefly discussed in footnote 8, the extra productivity due to a worker's starting level of general-purpose human capital were allowed to vary with job assignment. The reason is that, since this extra productivity would still increase with schooling, a sufficiently high value of β is enough to ensure that the wage increases with schooling, even after holding experience and job assignment fixed.

There are a number of other studies that find evidence consistent with schooling being positively related to promotions, and with schooling being positively related to the wage even after controlling for experience and job level. For example, similar to BGH, Medoff and Abraham (1980,1981) run Mincerian regressions that include job level and find that schooling is positively related to earnings even after one controls for experience and job level. They also study the factors that affect promotion and find that schooling is positively related to promotion probabilities. Similarly, McCue (1996) finds (using evidence from the PSID) that promotion probabilities are positively related to schooling level, although this finding is statistically significant only for white men; for other demographic groups, the coefficients typically are positive but are not statistically significant.

In a very recent study using German data, Lluís (2001) also looks at the effect of schooling on promotion rates. She first shows that schooling is significantly positively related to promotions. She then incorporates a lagged value for each worker's wage growth into her empirical analysis and shows that this weakens the correlation between schooling and promotions. Notice that, consistent with our discussion following Corollary 2, this is exactly what our model predicts should happen. The lagged value of a worker's wage growth is a noisy measure of the worker's ability to learn on the job, and our model thus predicts that the correlation between schooling and promotions should be weaker when lagged wage growth is included as an explanatory variable.

We now turn to the vast empirical literature concerning schooling and earnings that employs the earnings function first explored in Mincer (1974); see Willis (1986) for an early survey of this literature and Card (1999) for a more recent survey. As is well known, Mincer estimates a log wage equation:

$$(4) \quad \log y = \beta_1 + \beta_2 S + \beta_3 X + \beta_4 X^2 + e,$$

where y is earnings, S is the number of years of schooling, X is labor-market experience, and e is an error term. Mincer's finding that $\beta_2 > 0$ is consistent with our first prediction: initial compensation (i.e., the wage received when entering the labor market) is increasing in the schooling level. This finding is also consistent with our second prediction: the absolute (rather than percentage) return to an extra year of experience is increasing in schooling. The reason is that, given the log specification that Mincer studied, $\beta_2 > 0$ means that the percentage return to schooling is constant with experience, implying that the absolute return to schooling is increasing with experience.

But just because (4) is the specification that Mincer and others have studied, it is not necessarily the correct specification. For example, in his original study Mincer added to (4) an experience-schooling interaction term and found a significant negative coefficient on this term. This finding introduces the possibility that the absolute return to an extra year of experience is not positively related to the schooling level but rather is either independent or negatively related.

A recent study concerning schooling and wage rates that does not take the standard Mincerian approach is Farber and Gibbons (1996). (Altonji and Pierret (2001) is related but takes an approach more similar to Mincer's.) In testing a pure-learning model of wage dynamics, Farber and Gibbons estimate the wage-level (as opposed to log-wage) specification given in equation (5):

$$(5) \quad y = \beta_1 + \beta_2 S + \beta_3 X + \beta_4 X^2 + \beta_5 S \times X + \beta_6 U + \beta_7 U \times X + e,$$

where U is a variable correlated with ability but unobserved by employers (such as a test score). The pure-learning theory that Farber and Gibbons consider predicts $\beta_5 = 0$ and $\beta_7 > 0$: the absolute return to an extra year of schooling is independent of experience, whereas the absolute return to a higher value for U is increasing with experience.

Farber and Gibbons estimate regression (5) with and without the U and $U \times X$ terms. When they include controls for changes in the return to education over time, they find results consistent with their prediction that $\beta_5 = 0$, both when they include the U and $U \times X$ terms and when they do not. When the U and $U \times X$ terms are included, their finding that $\beta_5 = 0$ is consistent with our model.¹⁴ But when the U and $U \times X$ terms are not included, our model predicts $\beta_5 > 0$, so these Farber-Gibbons results are problematic for our theory. But Farber and Gibbons also find that when the controls for changes in the return to education over time are not included, then $\beta_5 > 0$ when the U and $U \times X$ terms are not included, as our theory predicts. It is possible that data limitations in the Farber-Gibbons sample created a strong relationship between a worker's schooling level and the date at which the worker entered their sample, so that the true effect of schooling on the return to experience is better estimated without controls for changes in the return to education over time.¹⁵ Clearly, additional empirical investigation of this issue would be helpful.¹⁶

¹⁴ The logic here is as follows. Our model assumes a correlation between schooling (S_i) and ability to learn on the job (β_i), captured by the term $B(S_i)$. In our model, this correlation with β_i is the only way that schooling might have a measured effect on the return to experience. Thus, if variables correlated with ability to learn on the job are included in the regression specification and there is no remaining correlation between schooling and ability once those variables are included, then the prediction of our model is what Farber and Gibbons found: $\beta_5 = 0$ and $\beta_7 > 0$.

In summary, although there is not much direct evidence concerning our theoretical predictions, on three of the predictions - schooling being positively related to promotions, the correlation between schooling and promotions being weaker when a noisy measure of a worker's ability to learn on the job is included, and schooling being related to the wage even after controlling for experience and job level - the available evidence is supportive. For the remaining prediction - schooling being positively related to the return to experience - there is a single study that is not supportive, but we do not feel the evidence there is definitive. (We take as given that the evidence is consistent with our prediction that schooling positively affects the starting wage.) Hence, all told, we feel the available evidence supports our theoretical model, but clearly more evidence is needed before reaching any final conclusions.

III. AN EXPLANATION OF COHORT EFFECTS

In this section we provide an explanation for why wage and promotion dynamics inside firms might exhibit cohort effects. We begin by discussing the recent empirical literature on cohort effects and providing a basic intuition that explains how task-specific human capital can produce cohort effects, even in the spot-market framework that we analyze. We then present a variant of the model in Section II and show that it exhibits cohort effects. Finally, we analyze some other possible implications of task-specific human capital in this model.

Our idea of task-specific human capital is closely related to occupation- and industry-specific human capital. In each case, human capital is specific to the nature of the work, not specific to the firm. Hence, when capital is accumulated, multiple firms value the capital, so most (or even all) of the value of the capital will be reflected in the worker's wage. The main difference between our idea of task-specific human capital and occupation- and industry-specific human capital is how the idea is applied. Our main

¹⁵ Their sample consisted of workers in the NLSY between the ages of 14 and 21 on January 1, 1979 and data for these workers between 1979 and 1991, where a worker typically entered the sample when he finished his schooling. Given this sample, there is likely a strong relationship between a worker's schooling level and the year when he enters the sample, i.e., those with high values for schooling will typically enter later. This causes two potential problems for finding the true impact that schooling has on the return to experience. First, there are likely to be few data points for many of the workers whose schooling levels are high. Second, because many of the workers with high schooling levels enter the sample later, some of the impact that education has on the return to experience might be picked up by the controls Farber and Gibbons included for changes in the return to education over time.

¹⁶ See Habermalz (2003) for a recent study using the Current Population Survey that allows for non-linearity concerning how the return to education varies with experience, and finds results that suggest education positively affects the return to experience over a substantial part of workers' careers.

focus is on how the presence of task-specific human capital affects the promotion process, whereas analyses of occupation- and industry-specific human capital typically ignore promotions.¹⁷

A) Empirical Evidence and Basic Argument

There are two well known recent studies of cohort effects: Baker, Gibbs, and Homstrom (1994b) and Beaudry and DiNardo (1991). BGH study the personnel records of managers in a single firm over a twenty-year period. Regarding cohort effects, they find that a cohort's average wage upon entering the firm is an important determinant of the cohort's average wage years after entry. Of course, this would not be surprising if a cohort's average wage upon entry reflected differences in the composition (and thus average productivity) of a cohort. But BGH find no evidence that composition differences in race, sex, age, or education can explain a cohort's average wage years after entry.

Beaudry and DiNardo find cohort effects in panel data covering a large cross-section of occupations and industries. Their initial finding is that the unemployment rate in the year a worker enters the firm affects wages years after entry. But Beaudry and DiNardo conduct further analyses that allow both the current unemployment rate and the lowest unemployment rate since a worker's entry year to affect current wages. They find that each of these three regressors (entry unemployment, current unemployment, and lowest unemployment) is significant on its own, but that the lowest unemployment rate since entry is the only significant regressor when all three are included. These Beaudry-DiNardo findings suggest alternative interpretations of the BGH cohort effect. For example, the apparent influence of a cohort's entry wage on its current wage might disappear if one controlled for the highest wage paid to any entering cohort since the first cohort entered.¹⁸

In Gibbons and Waldman (1999a), we argued that the spot-contracting framework we considered in that paper (and this one) is inconsistent with the cohort effects found by BGH, and with the range of

¹⁷ For a formal analysis of occupation-specific human capital see Weiss (1971), while Neal (1995) and Parent (2000) are recent empirical studies of industry-specific human capital. Consistent with our discussion, however, none of these papers considers the implications for the promotion process.

¹⁸ MacLeod and Malcomson (1993) put forth a model that can explain the cohort effects found by Beaudry and DiNardo. In their analysis, fixed-wage contracts are employed to induce efficient general investments by the firm (i.e., investments that are not specific to a particular worker). These contracts are such that the wage moves only when one of the outside options changes, so that the relationship will be terminated if the wage remains unchanged (even if it is efficient for the relationship to continue). Although such contracts can explain the Beaudry and DiNardo finding that the best labor-market conditions since the start of a worker's job have a significant positive effect on the worker's current wage, we are skeptical concerning the ability of such contracts to match other important findings in the empirical literature concerning careers.

findings by Beaudry and DiNardo. Our argument was based on Beaudry and DiNardo's interpretation of their own results in terms of long-term insurance contracts, akin to those analyzed by Harris and Holmstrom (1982). Since our 1999a model assumes wages are determined each period by competition between a worker's current employer and prospective employers (and all parties are risk neutral), long-term insurance contracts are neither necessary nor feasible, so we concluded that cohort effects were inconsistent with our model. In this section we show that we were incorrect in our previous conclusion: a competitive framework is potentially consistent with cohort effects, even after controlling for composition differences, because a cohort's prior experiences at a firm can have a significant effect on actual or perceived productivity. In the next subsection, we develop an argument along this line and show how it can explain the BGH finding that a cohort's entry wage is correlated with the cohort's wage years later.

There are two key aspects to this new model: i) the state of the economy can be either good or bad, and ii) human-capital acquisition is task-specific, so some of a worker's acquired human capital goes unutilized when a worker is promoted and is assigned to a new set of tasks. When the economy is in the bad state, a higher proportion of young workers enter the firm at the low-level job. Since human capital is task-specific, some of a worker's acquired human capital goes unutilized when she is promoted, so a cohort hired in the bad state has a low average wage years later. (This result holds even if all the workers get to the high-level job eventually, because they have less time there and so are less productive). To capture the BGH finding that low wages at entry predict low wages later, we require that the cohort's entry wage be low in the bad state. Given the two states are sufficiently different, this occurs in our model because the bad state has both more entry workers on the low-level job and a lower entry wage on the high-level job.¹⁹

One interesting aspect of this argument is that, although we explain the BGH finding that a cohort's entry wage is correlated with the cohort's wage years later, this is not because the entry wage itself matters. Instead, what matters is the proportion of workers who start at the low-level job, which affects the number and productivity of workers in the high-level job years later. Our explanation for the

¹⁹ Although we focus on the BGH finding that a cohort's entry wage is correlated with the cohort's average wage years later, we believe our basic argument can also capture the Beaudry and DiNardo findings. As indicated, they find that when entry unemployment, current unemployment, and lowest unemployment since entry are all included as explanatory variables in a wage regression, it is only the lowest unemployment rate since entry that matters. Our approach will capture this finding (or at least that the lowest unemployment rate since entry matters more) if the lowest unemployment rate since entry is more highly correlated with prior assignments to high-level jobs and thus more highly correlated with the amount and quality of the cohort's human capital in the current period.

BGH cohort effects is consistent with further evidence reported by BGH: although they find no evidence that composition differences in race, sex, age, and education across cohorts explain their cohort-effects finding, they do find that a higher proportion of new workers entered the firm at lower job levels during 1976-1985, which is when the entry wage was low. This is exactly what our explanation predicts: there is a low wage upon entry when more workers enter the firm in the lower job levels, and it is the low skill acquisition (or the lack of transferability of skills acquired in the low job) that affects wages years later.²⁰

In addition to the BGH study, there are a number of recent empirical studies that find results consistent with our argument. First, Kahn (2002) uses the NLSY to study the effect that the national unemployment rate in the year a worker graduates from college has on the worker's earnings over her career. Consistent with our theoretical model, Kahn finds that graduating in a bad year has large negative effects on wages, both in the year of graduation and much later in the career. Second, Devereux (2002a) uses the PSID to investigate the effects that taking a good job (*i.e.*, a high-paying job) has on wages years later. (An important difference between this study and Beaudry and DiNardo's is that Devereux does not exclude workers who subsequently switch employers.) Also consistent with our theoretical approach, Devereux finds that there is a significant positive effect on wages received four years later, both for workers who remain with the same employer and for those who switch. Third, a related study that supports our general approach is Devereux (2002b), which shows that the average education level of new hires within occupations is higher when the economy is contracting rather than expanding. This is consistent with our model in that, for both the low-level and high-level jobs, the average education level for new hires is higher when the state of the economy is bad rather than good.²¹

²⁰ Note that what we call state of the economy could just as well be the state of the industry (or even state of the firm, if we were to introduce some firm-specific human capital into the model). These two reinterpretations are quite plausible for the financial services firm that BGH study.

²¹ Another related study is Welch (1979), which considers the effects of cohort size on earnings. Welch finds that being a member of a large cohort has a significant negative effect on the worker's wage when she enters the labor force and a smaller, but still negative effect, later in the worker's career. He interprets this as limited substitutability in production across cohorts, where this limited substitutability gets less important as the cohort ages. An alternative interpretation is that task-specific human capital is important. In this interpretation, when a large cohort enters the labor market, because there are a limited number of high-level positions available to entry-level workers, the cohort earns on average lower wages and works on average in lower-level jobs. In turn, if human capital is task specific, then there will be a subsequent negative effect on wages as the cohort ages. The reason is that, if a large cohort enters the labor force on average in lower-level jobs, then the presence of task-specific human capital lowers subsequent wages because the human capital the cohort acquires while young is less valuable after promotions take place.

As a final point, although our analysis focuses on how task-specific human capital can explain cohort effects, the basic point is more general: cohort effects can arise whenever a cohort's past experiences affect subsequent actual and/or perceived productivity. For example, suppose there is no task-specific human capital, but the rate at which workers acquire human capital depends on how "busy" the firm is (e.g., workers learn more from "interesting" projects, but such projects are scarce in slack times). Then a cohort that enters the firm when it is not busy may receive a low wage upon entry and also do poorly years later because little was learned during the entry period. Similarly, suppose there is no task-specific human capital and how busy a firm is does not matter, but asymmetric learning causes a worker's job-assignment history to serve as a signal of ability (as initially explored in Waldman (1984b,1990) and Ricart i Costa (1987)). That is, suppose more workers enter the labor market at low-level jobs when the state of the economy is bad, and promotions are rare as long as the economy stays bad. Then, if prospective employers interpret a worker's job-assignment history as a signal of the worker's ability, but fail to account for the state of the world when a worker was young, then a worker's cohort can affect the worker's compensation years later.

B) Model

There are three main differences between the model explored here and the one analyzed in Section II. The first difference concerns the idea that human capital is now task-specific. But to keep the model simple, we restrict this effect to job 2. That is, we assume that job 1 is analogous to both jobs in Section II (and all the jobs in our 1999a paper): past labor-market experience improves a worker's current productivity on job 1, and the extent of this productivity improvement is independent of how much of the past experience was on which jobs. Formally, we assume that if worker i has experience x_{it} in period t , then this worker's output in job 1 in that period is

$$(6) \quad y_{i1t} = d_1 + c_1 [f(x_{it}) + \epsilon_{i1t}],$$

similar to equation (2) in Section II. For job 2, however, we make a different assumption. Now skill acquisition is task-specific: past experience improves a worker's current productivity on job 2, but experience on job 2 creates a larger productivity improvement than does experience on job 1. Formally, we assume that if worker i has experience x_{it} in period t then this worker's output in job 2 in that period is

$$(7) \quad y_{i2t} = d_2 + c_2 [f(x_{i2t} + x_{i1t}) + \epsilon_{i2t}],$$

where x_{ijt} is the past experience that worker i has accumulated in job j prior to period t , so $x_{i1t} + x_{i2t} = x_{it}$, and $(0,1)$ is a parameter that reflects the extent to which experience on job 1 creates a productivity improvement on job 2. That is, because acquired human capital is task-specific in job 2, when a worker is promoted from job 1 to job 2, a previous period that the worker spent in job 1 counts as only θ of a period in the skill-acquisition function.²² Although equations (6) and (7) no longer include the productivity term $G(S_i)$ since it is not needed for any of the results below, we continue to assume heterogeneity in schooling, because of the role it plays in initial job assignments.

The second difference between this section's model and Section II's is that now there are two states of the world, where the state of the world affects relative productivities and hence job assignments. The simplest way to capture this possibility is to have the state of the world affect the productivity of the high-level job.²³ In particular, we assume that there are two intercepts for the high-level job. If the economy is in the good state then $d_2 = d_2^G$; if it is in the bad state then $d_2 = d_2^B$, where $d_2^G > d_2^B$. We further assume that the probability that the economy is in the good state in any period is q , independent across periods, and that the realization of the state is determined and publicly observed at the beginning of each period (before wages and job assignments are determined). When the economy is in the good state, holding all else equal, more workers will be assigned to the high-level job, because $d_2^G > d_2^B$.

The third difference between this section's model and Section II's is that, because we want to explore how the history of job assignments for young workers affects job assignments and wages when workers become old, we no longer restrict the analysis to parameterizations for which all entry workers are assigned to the low-level job. To the contrary, we now restrict the analysis to parameterizations where some entry workers are assigned to the high-level job. In particular, we assume that a worker with the highest schooling level (S_N) is assigned to the high-level job in the worker's first period in the labor market if the state of the world is good, and that a worker with the lowest schooling level (S_1) is assigned to the low-level job in the worker's first period in the labor market if the state of the world is bad.²⁴

²² As indicated, to keep the model simple, we have allowed task-specific human capital only in job 2. We could also allow task-specific human capital in job 1. If we then further restricted the analysis to parameterizations for which there are few (if any) demotions (which is what most of the data shows), then this alternative specification would produce basically the same results as we derive.

²³ We could assume that the state of the world affects productivity in both jobs. All of our qualitative results would still go through as long as the effect that the state of the world has on the high-level job is larger than its effect on the low-level job.

Although the model in this section seems closely related to the model in Section II and the main model of our 1999a paper, the model in this section is in fact much more conceptually challenging. In Section II and our earlier work, a worker's output in any period was independent of the worker's job-assignment history. As a result, in each period, the assignment of workers to jobs was simple: a worker is always assigned to the job that maximizes that period's expected output. But that assignment rule is not optimal here. Instead, the current job assignment is now partly an investment in the worker's expected future productivity, so the worker's optimal assignment may not maximize current productivity. Because of this, to keep the intuition behind cohort effects clear, we begin by assuming that careers last two periods (i.e., $T=2$). Note that for this part of the analysis we use the terms "young" and "old" to refer to workers in their first and second periods in the labor market, respectively.²⁵ We then consider what happens if careers last for more than two periods.

C) Analysis

As indicated, we start by considering the case $T=2$. We begin our analysis of this case by considering how workers are assigned to jobs. Consider first what happens when workers are old. Since there are no subsequent productivities to be concerned with, when a worker is old, she will be assigned to the job that maximizes her expected output that period. But the solution to this second-period assignment problem depends on the worker's job assignment when young. In particular, because some human capital goes unutilized when a worker is promoted, more old workers will be assigned to the high-level job this period if more young workers were assigned to the high-level job last period.

Now consider assignments when workers are young. In contrast to what happens when workers are old, the optimal job assignments when workers are young do not simply maximize expected productivity in the current period. Rather, the assignments now take into account the effect that this period's assignment has on next period's optimized productivity. In particular, because assigning a young worker to the high-level job increases the worker's output in the following period if she is again assigned

²⁴ We also restrict the analysis to parameterizations such that at least one schooling group is assigned to the high-level job in the good state of the world and to the low-level job in the bad state.

²⁵ In this case we assume that β is sufficiently close to one and/or $f(1)$ is sufficiently small that some promotions occur when a worker becomes old. If β were sufficiently small and $f(1)$ were sufficiently large, there would be no promotions in equilibrium because the underutilization of acquired human capital associated with a promotion would make all promotions unprofitable.

to the high-level job, assignments will be biased towards the high-level job, relative to assignments that optimize today's productivity. The other important aspect of assignments (both when workers are young and when they are old) is that more workers are assigned to the high-level job in the good state of the world than in the bad state. This follows because the high-level job is even more productive relative to the low-level job in the good state of the world.

We formalize the above discussion in Proposition 2. Let γ_{it}^e denote the critical value of γ_{it}^e such that young worker i with schooling level S in period t is assigned to the high-level job if $\gamma_{it}^e > \gamma_{it}^K(S)$ and is assigned to the low-level job if $\gamma_{it}^e < \gamma_{it}^K(S)$, where $K=G,B$ is the state of the world in period t . Similarly, let $\omega_{it}^{K,J}(S)$ denote the critical value of ω_{it}^e such that old worker i with schooling level S in period t is assigned to the high-level job when $\omega_{it}^e > \omega_{it}^{K,J}(S)$ and is assigned to the low-level job when $\omega_{it}^e < \omega_{it}^{K,J}(S)$, where $K=G,B$ is again the state of the world in period t and $J=1,2$ denotes the worker's job assignment in period $t-1$. Finally, let γ_Y^{K*} denote the critical value of γ_{it}^e if the assignment of a young worker in state K were to maximize current productivity, and let $\omega_O^{K,J*}$ denote the critical value of ω_{it}^e if the assignment of an old worker in state K were to maximize current productivity, given that the worker was assigned to job-level J in the previous period.

Proposition 2: If $T=2$, then job assignments satisfy i) through iv).

- i) $\gamma_Y^K(S) < \gamma_Y^{K*}$ for all S and $K=G,B$.
- ii) $\omega_O^{K,J}(S) = \omega_O^{K,J*}$ for all S and all (K,J) pairs, $K=G,B$ and $J=1,2$.
- iii) $\gamma_Y^G(S) < \gamma_Y^B(S)$ for all S and $\omega_O^{G,J}(S) < \omega_O^{B,J}(S)$ for all S and $J=1,2$.
- iv) $\omega_O^{K,H*} < \omega_O^{K,L*}$ for $K=G,B$.

We now turn our attention to cohort effects. Let W_Y^K denote the average wage paid to young workers in state of the world K , and $W_O^{K,K}$ the average wage paid to old workers when the current state of the world is K and the state of the world in the previous period was K .

Proposition 3: Suppose $T=2$. Then holding all the other parameters fixed, there exists a value d_2^G such that, if $d_2^G > d_2^B$, then there is a unique equilibrium that satisfies i) and ii).

- i) $W_Y^G > W_Y^B$.

ii) $W_0^{K,G} > W_0^{K,B}$ for $K=G,B$.

Proposition 3 tells us that, if d_2^G is sufficiently large, then the model exhibits a cohort effect like that found by BGH: a cohort that enters a firm with a high entry wage will earn more when the cohort becomes old. To see this consider first what happens when a cohort becomes old. Holding constant this period's state of the world, the old cohort's average pay will be higher when the previous period's state was good rather than bad. This result occurs because more young workers are assigned to the high-level job in the good state, and being assigned to the high-level job when young increases productivity in the high-level job when old.

Now consider what happens when a cohort enters the labor market. As stated earlier, relative to entry in the bad state, more workers will be assigned to the high-level job in the good state. In turn, there are two factors affecting whether the average wage for young workers is higher in the good state or the bad state. First, because the high-level job is more productive in the good state, the average wage is higher in the good state since those who would be assigned to the high-level job in both states are paid more in the good state.

Second, again because the high-level job is more productive in the good state, some workers are assigned to the high-level job only in the good state and these workers may earn on average either more or less in the good state than in the bad state. They can earn more because the high-level job is more productive in the good state. But there is also the possibility they earn less because a young worker is willing to work in the high-level job for less given that human capital accumulated on that job is on average more valuable when the worker becomes old. In other words, because of the future return associated with working in the high-level job when young, a young worker's wage can in fact fall when the state of the world is good rather than bad and the worker's job assignment switches from low to high. The assumption d_2^G sufficiently large ensures that the first of these factors is the dominant one, and thus that the model exhibits cohort effects in the sense that a good state of the world translates into high wages both when the cohort is young and when it is old.

As discussed earlier, one interesting aspect of our explanation for the BGH finding is that, although a cohort's average wage when young is correlated with its average wage when old, the wage when young is only a proxy for the cohort's job assignments when young (and consequently for the

amount and quality of human capital that the cohort acquires when it is young). This is important for two reasons. First, from a competitive standpoint, it would be surprising for the average wage a cohort earns when young to affect the average wage the cohort earns when it ages, but it is not surprising for the amount and quality of the human capital that a cohort accumulates when it is young to affect the cohort's average wage when it ages. Second, as discussed earlier, this aspect of our explanation for the BGH cohort-effects finding is consistent with another aspect of their empirical analysis: in their data, when the entry wage was low, more workers entered the firm at lower job levels.

We now consider what happens when $T > 2$. In this analysis we consider both cohort effects and other aspects of wage and promotion dynamics when there is job-specific human capital.²⁶

Proposition 4: If $T > 2$, then i) through iii) are satisfied.

- i) Holding all other parameters fixed, there exists a value d_2^G such that, if $d_2^G > d_2^G$, then there is a positive correlation between a cohort's average wage when young and its expected compensation over the last $T-1$ periods the cohort is in the labor market.
- ii) Hold all other parameters fixed other than β and let β be the highest value for β such that $d_1 + c_1 f(T-1) > d_2^G + c_2 f(\beta(T-1))$ for all $\beta \in [L+B(0), H+B(N)]$ (such a value will necessarily exist if $f(0)$ is sufficiently small). If $\beta < \beta$, then there exists a smallest value t , $t \in [1, T-1]$, such that any worker who spends her first t periods in job 1 is not promoted in any of her last $T-t$ periods in the labor market.
- iii) There are parameterizations in which some promotions are associated with wage decreases.

Proposition 4 tells us that there are three results that characterize the case $T > 2$.²⁷ The first captured in i) is that, not surprisingly, when d_2^G is sufficiently large the model continues to exhibit cohort effects: the average wage paid to a cohort in its first period in the labor market is correlated with average wages paid to the cohort later. As above, this occurs because acquired human capital is task-specific, so

²⁶ As for our analysis of $T=2$, we continue to restrict the analysis to parameterizations such that there are some promotions in equilibrium (see footnote 25).

²⁷ When workers are in the labor market for more than two periods, this model is similar to the analysis in Gibbons and Waldman (1999a). Thus, as was true for the model of Section II, this model should be consistent with a variety of findings in BGH and elsewhere, such as that wage increases are serially correlated, real-wage decreases are not rare but demotions are, and workers who receive large wage increases early in their stay at one level of the job ladder are promoted quickly to the next. Here we focus on results other than those found in our 1999a analysis.

the state of the world in the worker's first period in the labor market influences not only job assignments and wages in that first period, but consequently also job assignments and wages later in these careers.

The second result captured in ii) is that, if the specificity of human capital is sufficiently large, i.e., β is sufficiently small, then a worker who spends substantial time in job 1 at the beginning of her career can get stuck in that job with no possibility of a subsequent promotion. That is, it can be the case that a worker who spends her first t ($<T$) periods in the low-level job has a zero probability of being promoted to the high-level job in any of the last $T-t$ periods. There are two factors leading to this result. First, as the number of periods in job 1 grows, the loss in lifetime productivity upon promotion becomes large (because some job-1 human capital will be underutilized after promotion). Second, again as the number of periods in job 1 grows, the gain in lifetime productivity upon promotion from human-capital acquisition on job 2 becomes small, because fewer periods remain in the worker's career.

This second result is consistent with BGH evidence concerning promotion rates for workers in level 2 as a function of the number of years the worker has been in level 2 or below. Extending the logic from the above discussion to the case of three job levels yields that a worker who spent many years in level 2 or below would have little or no probability of further promotion. The logic is that the large number of years in level 2 or below means that much human capital would go unutilized with a promotion, and fewer years would exist after a promotion to build human capital. BGH find just such a result: the promotion rate is less than two percent for workers who were at level 2 or below for at least fifteen years (compared to an overall promotion rate for workers at level 2 of nineteen percent).²⁸

The last result captured in iii) is that, in contrast to what was true in Gibbons and Waldman (1999a), there are parameterizations in which some (possibly many) promotions will be associated with wage decreases. This also follows from the presence of task-specific human capital. Because being in the high-level job increases productivity in the high-level job in subsequent periods, in period t a worker will sometimes prefer a lower wage on the high-level job to a higher wage on the low-level job.

Extending this argument to incorporate the worker's wage on the low-level job in period $t-1$ yields that, if

²⁸ Our general approach also provides a second potential explanation for this finding: workers who work many years in each of the two lowest job levels may have sufficiently low ability to learn on the job that it is rarely optimal to promote such workers to level 3.

task-specific human capital is significant (i.e., β is small), some promotions may be associated with wage decreases.²⁹

Notice that this argument is similar to a classic argument concerning wage profiles and interfirm mobility over the career (e.g., Topel and Ward (1992)). In this argument, there are no promotions and no turnover, but firms vary in terms of the rate at which a worker in the firm accumulates human capital over the worker's career. As a result, young workers are willing to work at high-capital-accumulation firms for less because of the effect on subsequent productivity and subsequent wages. The result is that firms with fast accumulation of human capital have lower starting wages but higher wage growth over the career. Our logic here is the same except that fast accumulation of human capital is associated with promotions, so some promotions are likely to be associated with wage decreases.

Given the above, the obvious question is, to what extent are some promotions associated with wage decreases even though on average promotions are associated with wage increases? BGH do not report the percentages of their promotions associated with wage increases and decreases. However, as discussed earlier, McCue (1996) also studies promotions and wage changes and she does report information concerning the percentages of promotions associated with wage increases and decreases. Consistent with BGH and many other studies, McCue finds that promotions on average are associated with large wage increases. But she also finds that roughly twenty-five percent of promoted workers in her sample received real-wage decreases.³⁰ This finding is inconsistent with our 1999a paper and the model of Section II, but is consistent with a world where task-specific human capital is important.

IV. LIMITATIONS OF OUR APPROACH

We feel that in our 1999a paper and here we have made a strong case that our approach is able to capture a variety of facts concerning wage and promotion dynamics in a large class of settings. But there

²⁹ In our model, a worker's wage each period equals that period's expected productivity, so the idea that promotions are sometimes associated with wage decreases means the model also predicts that promotions will sometimes be associated with productivity decreases. Task-specific human capital, therefore, is an additional explanation to the ones identified by Lazear (2002) for why productivity might sometimes fall after a promotion occurs. Further, the fact that promotions can be associated with wage decreases means that, in addition to the explanation developed in Section II concerning schooling, task-specific human capital serves as an explanation for why adjacent job levels can have overlapping wage distributions.

³⁰ McCue reports the real-wage increase for promoted workers at the twenty-fifth percentile broken down by race and gender. She finds that for white men this value is negative but small while for the other categories - white women, black men, and black women - the value is positive but small.

are some settings in which we believe our approach does not apply. In this section we discuss limitations of our approach.

There are two key interrelated aspects of our approach that are important for generating predictions, but that also limit its applicability. First, there is only spot-market contracting; there are no insurance contracts. Second, there is no firm-specific human capital (and the value of task-specific human capital does not vary across employers). In combination, these two assumptions imply that wages in every period are determined by the wage offers of prospective employers. In a setting in which each level of a firm's job ladder exhibits a significant proportion of workers both entering and exiting the firm (as is true for the BGH firm), we believe that these aspects of our approach make sense and help the model better match the data. The fact that each level of the job ladder exhibits significant entry and exit suggests that firm-specific human capital is limited. Further, since there is a significant probability of a worker leaving at every job level, the wage offers of prospective employers are likely to be important in determining what a worker's initial employer is willing to pay.³¹

But there are other settings in which our approach is probably not a good one, exactly because of the two assumptions mentioned above. For example, consider a firm with an internal labor market like that described by Doeringer and Piore (1971). Among other aspects, this firm is characterized by ports of entry (meaning little or no entry at higher levels of the job ladder) and little exit at higher levels of the job ladder. The fact that there is little turnover at higher levels of the job ladder suggests that the labor-market practices of this firm may not be driven by the wage offers of prospective employers. Rather, it may be that significant firm-specific capital insulates the firm's workers from the outside market, so some alternative theory of wage and promotion dynamics is needed to understand careers in this firm.

Our approach shares one feature with the pure-form internal labor market (ILM) described by Doeringer and Piore: an important role for job levels in wage determination. But in the pure-form ILM, wages are attached to jobs. (That is, there is little or no wage variation within a job, except perhaps formulaic variation with seniority.) In our approach, in contrast, a worker's wage is determined by her

³¹ We feel that turnover in the data makes it more likely that our model is applicable, even though our model itself exhibits no turnover. In our model, in each period a worker is close to indifferent between moving and staying (one interpretation of our model is that there is an infinitesimal moving cost that results in no turnover, but means the worker prefers staying only by an infinitesimal amount). It would be easy to make small changes in the model that result in some turnover but cause few other changes.

productivity. Because our model incorporates learning and heterogeneity in human-capital acquisition, wage variation within a job cannot be explained solely by variation in seniority. Thus, within-job wage variation may be another way (in addition to entry and exits at higher levels of the job ladder) to identify settings in which our approach applies rather than the pure-form ILM described by Doeringer and Piore.

V. CONCLUSION

A large number of theories have been put forth to explain wage and promotion dynamics inside firms. We believe that an important goal in such model-building should be to explain broad patterns of evidence rather than isolated stylized facts. In our 1999a paper, we showed that a model that combines job assignment, human-capital acquisition, and learning captures a number of findings concerning wage and promotion dynamics found in BGH and elsewhere, and that our theory matches the evidence better than available alternatives do. In this paper we extend that earlier analysis in two ways. First, we incorporate schooling into the framework and develop several predictions, most of which match well with available evidence concerning the role of schooling in wage and promotion dynamics. Second, we show that the introduction of task-specific human capital allows the model to capture BGH's finding of cohort effects, which in our earlier paper we incorrectly argued was inconsistent with our theoretical approach.

The other main message of this paper is the potential significance of task-specific human capital. Beginning with the seminal work of Becker, most of the literature concerning human capital has focused on the concepts of general-purpose and firm-specific human capital. We believe, however, that task-specific human capital is potentially as prevalent and important as are general-purpose and firm-specific human capital. Clearly, much of human-capital acquisition involves becoming more proficient at the task or tasks being performed. Thus, rather than human capital going unutilized when a worker switches firms (as is the case when human capital is firm-specific), human capital goes unutilized when a worker switches jobs and is assigned to a new set of tasks, whether the switch entails staying within the same firm or moving across firms. In this sense, our notion of task-specific human capital is akin to occupation- or even industry-specific human capital. In this paper we have begun to explore some of the implications of task-specific human capital. We showed that this concept can explain cohort effects, but we conjecture that there are a variety of other applications such as job design, job assignment, labor mobility, labor demand, and even business strategy.

There are a number of directions in which the analysis in this paper could be extended, including the following three. First, as just indicated, one could extend our analysis of task-specific human capital to a variety of topics. Second, one could incorporate into our framework other modeling elements that have received significant theoretical attention such as incentives and asymmetric learning, to see whether the addition of such elements makes the model a better match with the evidence.³² Third, one could integrate our focus on careers inside firms with the extensive theoretical literature that attempts to explain evidence concerning wage dynamics and mobility between firms (see, e.g., Burdett (1978), Mortensen (1978), and Jovanovic (1979)). Since promotions and turnover are alternative exit routes from a worker's current job, there should be a single theoretical framework that captures the empirical evidence on both topics.³³

APPENDIX

In the Appendix we present proofs of the Propositions and Corollaries.

Proof of Proposition 1: This proof is similar to the proof of Proposition 2 in Gibbons and Waldman (1999a). Because learning is symmetric and there is no firm-specific human capital, competition among firms each period yields both efficient job assignment and wages equal to expected output. Given this, we compute worker i 's expected on-the-job human capital in period t , h_{it}^e , in (3) and then the worker's expected output in job j as $Ey_{ijt} = d_j + G(S_i) + c_j h_{it}^e$. Note that the linearity of the production function in (2) is key here: without linearity, expected output would not equal the output of a worker known to have on-the-job human capital equal to h_{it}^e . We now have that, given efficient job assignment and wages equal to expected output, if $h_{it}^e < h_{it}^e$ then worker i is assigned to job 1 in period t and paid $w_{it} = d_1 + G(S_i) + c_1 h_{it}^e$, while if $h_{it}^e > h_{it}^e$ then worker i is assigned to job 2 in period t and paid $w_{it} = d_2 + G(S_i) + c_2 h_{it}^e$ (see footnote 9). Note here the fact (2) reads $d_j + G(S_i) + c_j(h_{it} + \epsilon_{ijt})$ means that the signal about ability that can be extracted from output does not vary in its signal-to-noise ratio as a function of j . Thus, there is no way to use job assignment to change the speed of learning about ability, so job assignment is determined by current

³² One prominent approach to incentives in firms involves tournaments. Papers concerning the tournament issue include Lazear and Rosen (1981) and Rosen (1986), while asymmetric learning is studied in Waldman (1984b, 1990), Greenwald (1986), and Gibbons and Katz (1991) among others. See Zaboynik and Bernhardt (2001) for an interesting recent paper that combines the tournament and asymmetric-learning approaches.

³³ See Ghosh (2002) for a start in this direction. Also, Munasinghe (2001) explores a model that captures a number of findings concerning wages and turnover, although she does not explore the extent to which her model captures findings concerning careers inside firms.

productive efficiency (i.e., maximizing expected output this period), which in turn is solely a function of the worker's current expected on-the-job human capital. Q.E.D.

Proof of Corollary 1: Given Proposition 1 and that the analysis is restricted to parameterizations such that it is efficient for each worker to be assigned to job 1 in the first period of the worker's career (i.e., $[p_H + (1-p_L) + B(N)]f(0) < \dots$), the wage of worker i in her first period in the labor market can be written as $d_1 + G(S_i) + c_1[p_H + (1-p_L) + B(S_i)]f(0)$. Given $G > 0$ and $B > 0$, we immediately have that the worker's wage is a strictly increasing function of the worker's schooling level. Q.E.D.

Proof of Corollary 2: Because the signal z_{it} is independent of job assignment and schooling level, the amount in any period that has been learned about θ_i is independent of a worker's schooling level and past job assignments. In turn, since a worker is assigned to the high-level job whenever $[z_{it} + B(S)]f(x_{it}) > \dots$ and $B > 0$, we immediately have that for any (X, S) pair, where $1 \leq X \leq T-1$ and $S \leq N-1$, $(X, S+1) \succ (X, S)$. In turn, since by assumption $[z_{it} + B(S)]f(T-1) < \dots < [z_{it} + B(S)]f(T-1)$ for all S , we have that this inequality must be strict if X is sufficiently close to $T-1$.

Proof of Corollary 3: Consider first values for X and $X+1$ such that all workers of those experience levels are assigned to job 1. Let $\theta_i^* = p_H + (1-p_L)$. Given expectations each period must be correct on average and output in job 1 is linear in on-the-job human capital, we have that $W(X, S)$ is given by (A1).

$$(A1) \quad W(X, S) = d_1 + c_1[\theta_i^* + B(S)]f(X)$$

In turn, the return to experience for a fixed schooling level is given by (A2).

$$(A2) \quad W(X+1, S) - W(X, S) = c_1[\theta_i^* + B(S)][f(X+1) - f(X)]$$

Given $B > 0$ and $f' > 0$, (A2) immediately yields that $W(X+1, S) - W(X, S)$ is strictly increasing in S . Note that the increase in the return to an extra year of experience as schooling is increased from S to $S+1$ is given by (A3).

$$(A3) \quad [B(S+1) - B(S)]c_1[f(X+1) - f(X)]$$

Now consider values X and $X+1$ such that some workers may be assigned to job 2 but there are no demotions. Consider a worker with experience level X and schooling level S who is assigned to job 1 and for whom there is an expectation concerning θ_i given by θ_i^* . Given expectations each period must be correct on average and output in job 1 is linear in on-the-job human capital, if the worker is assigned to job 1 with probability one next period, then her expected wage increase is given by $c_1[\theta_i^* + B(S)][f(X+1) - f(X)]$. Similarly, consider a worker with experience level X and schooling level S assigned to job 2 for whom there is an expectation concerning θ_i given by θ_i^* . This worker's expected wage increase is given by $c_2[\theta_i^* + B(S)][f(X+1) - f(X)]$. Now consider a worker with experience level X and schooling level S initially assigned to job 1 for whom there is an expectation concerning θ_i given by θ_i^* , where there is a strictly positive probability the worker will be assigned to job 2 in the following period. Because

expectations are correct on average every period, the expected wage increase is given by $rc_1[\beta + B(S)][f(X+1)-f(X)]+(1-r)c_2[\beta + B(S)][f(X+1)-f(X)]$ where $0 < r < 1$. These expressions yield that integrating the wage increases over all workers of experience X and schooling S yields (A4), where $0 < R < 1$.

$$(A4) \quad W(X+1,S)-W(X,S)=Rc_1[\beta + B(S)][f(X+1)-f(X)]+(1-R)c_2[\beta + B(S)][f(X+1)-f(X)]$$

Now consider workers of the same experience level but schooling level $S+1$. Because a worker whose initial expectation for w is given by β may now be assigned to job 2 in either period, this worker's expected wage increase must be greater than or equal to $c_1[\beta + B(S+1)][f(X+1)-f(X)]$. Similarly, a worker whose initial expectation for w is given by β must now be assigned to job 2 in each period, so this worker's expected wage increase is $c_2[\beta + B(S+1)][f(X+1)-f(X)]$. Now consider a worker whose initial expectation for w is β . Because for any realization for the stochastic term in which the worker is assigned to job 2 given $X+1$ and S the worker is also assigned to job 2 given $X+1$ and $S+1$, this worker's expected wage increase must be greater than or equal to $rc_1[\beta + B(S+1)][f(X+1)-f(X)]+(1-r)c_2[\beta + B(S+1)][f(X+1)-f(X)]$. Given these expressions, integrating the wage increases over all workers of experience level X and schooling level $S+1$ yields (A5).

$$(A5) \quad W(X+1,S+1)-W(X,S+1) \geq Rc_1[\beta + B(S+1)][f(X+1)-f(X)]+(1-R)c_2[\beta + B(S+1)][f(X+1)-f(X)]$$

We now have that the increase in the return to an extra period of experience due to the schooling level increasing from S to $S+1$ is greater than or equal to the expression given in (A6).

$$(A6) \quad [B(S+1)-B(S)][Rc_1+(1-R)c_2][f(X+1)-f(X)]$$

Together, (A3) and (A6) yield that, given no demotions, $W(X+1,S)-W(X,S)$ is strictly increasing in S .

We now consider what happens when demotions are possible. Consider values for X and $X+1$ such that workers of those experience levels can be assigned to either job and there are some demotions as experience increases from X to $X+1$. Consider a worker with experience level X and schooling level S who is assigned to job 1 and for whom there is an expectation concerning w given by β . As before, if the worker is assigned to job 1 with probability one next period, then her expected wage increase is given by $c_1[\beta + B(S)][f(X+1)-f(X)]$. Consider a worker with experience level X and schooling level S assigned to job 2 for whom there is an expectation concerning w given by β , where there is a positive probability the worker will be assigned to job 1 in the following period. This worker's expected wage increase is strictly less than $c_2[\beta + B(S)][f(X+1)-f(X)]+mc_2 \beta f(X+1)$, where m is the probability of demotion and β is the absolute value of the average reduction in the expectation concerning w when the worker is demoted.

Consider a worker with experience level X and schooling level S assigned to job 1 for whom there is an expectation concerning w given by β , where there is a strictly positive probability the worker will be assigned to job 2 in the following period. As before, the expected wage increase for this worker is given by $rc_1[\beta + B(S)][f(X+1)-f(X)]+(1-r)c_2[\beta + B(S)][f(X+1)-f(X)]$, where $0 < r < 1$. These expressions yield that integrating the wage increases over all workers of experience X and schooling S yields (A7), where M is the overall probability of demotion for these workers and $0 < R < 1$.

$$(A7) \quad W(X+1,S)-W(X,S) < R c_1 [\theta + B(S)][f(X+1)-f(X)] + (1-R)c_2 [\theta + B(S)][f(X+1)-f(X)] + M c_2 [\theta - L] f(X+1)$$

Now consider workers of the same experience level but schooling level $S+1$. Because a worker whose initial expectation for θ is given by θ may now be assigned to job 2 in either period, this worker's expected wage increase must be greater than or equal to $c_1 [\theta + B(S+1)][f(X+1)-f(X)]$. Because a worker whose initial expectation for θ is θ may be assigned to job 1 in the following period, this worker's expected wage increase must be greater than or equal to $c_2 [\theta + B(S+1)][f(X+1)-f(X)]$. Because a worker whose initial expectation for θ is θ has a higher probability of being assigned to job 2 in the following period when the schooling level is $S+1$ rather than S , this worker's expected wage increase must be greater than or equal to $r c_1 [\theta + B(S+1)][f(X+1)-f(X)] + (1-r)c_2 [\theta + B(S+1)][f(X+1)-f(X)]$. These expressions yield that integrating the wage increases over all workers of experience X and schooling $S+1$ yields (A8).

$$(A8) \quad W(X+1,S)-W(X,S) > R c_1 [\theta + B(S+1)][f(X+1)-f(X)] + (1-R)c_2 [\theta + B(S+1)][f(X+1)-f(X)]$$

Comparing (A7) and (A8) yields that for any parameterization, if demotions are sufficiently rare, i.e., M is sufficiently small, then $W(X+1,S)-W(X,S)$ is strictly increasing in S . Q.E.D.

Proof of Corollary 4: Let $W_j(X,S)$ be the average wage paid to workers of experience level X and schooling level S who are assigned to job j given a value for θ . Because the productivity due to a worker's starting level of general-purpose human capital, $g(S)$, does not depend on the worker's job assignment, varying θ will not affect job assignments. Hence, $W_j(X,S)$ can be written as $W_j^0(X,S) + g(S)$. This, in turn, yields that for all $0 \leq X \leq T-1$, $S \leq N-1$, and $j=H,L$, $W_j(X,S+1)-W_j(X,S) = W_j^0(X,S+1) - W_j^0(X,S) + [g(S+1)-g(S)]$. Given $g' > 0$, this expression is strictly positive given θ sufficiently large. Q.E.D.

Proof of Proposition 2: Consider first an old worker with schooling level S . Since this is the last period of the worker's career, competition across firms means the worker will be assigned to the job that maximizes productivity that period and paid a wage equal to that productivity. We also know that, since $c_2 > c_1$, if in state of the world K an old worker who worked in job J last period with expected ability to learn on the job equal to θ is efficiently assigned to job 2, then in state of the world K an old worker who worked in job J last period with expected ability to learn on the job equal to θ , $\theta > \theta^*$, must also efficiently be assigned to job 2. This means that for each S and each (K,J) pair, $K=G,B$ and $J=1,2$, there exists a value $\theta^{K,J}(S)$ such that if period t is characterized by state of the world K then old worker i previously assigned to job J is assigned to job 2 if $\theta_i^e \geq \theta^{K,J}(S)$ and is assigned to job 1 if $\theta_i^e < \theta^{K,J}(S)$. Further, since assignments maximize that period's expected output, we also have $\theta^{K,J}(S) = \theta^{K,J*}$ for all S and (K,J) pairs, $K=G,B$ and $J=1,2$. This proves ii).

Consider again $\theta^{K,2*}$. This is the value for θ_i^e such that in state of the world K an old worker who was in job J in the previous period has the same expected output in each of the two jobs. In other words, $\theta^{K,2*}$ satisfies (A9) while $\theta^{K,1*}$ satisfies (A10).

$$(A9) \quad d_1+c_1[o^{K,2*}f(1)]=d_2^K+c_2[o^{K,2*}f(1)]$$

$$(A10) \quad d_1+c_1[o^{K,1*}f(1)]=d_2^K+c_2[o^{K,1*}f(1)]$$

(A9) yields $o^{K,2*}=(d_1-d_2^K)/[c_2f(1)-c_1f(1)]$ while (A10) yields $o^{K,1*}=(d_1-d_2^K)/[c_2f(1)-c_1f(1)]$. A comparison of these expressions yields $o^{K,2*} < o^{K,1*}$ for $K=G,B$. Note, if $o^{K,2*} < o^{K,1*}$ for $K=G,B$ or $H+B(N)$ $o^{K,2*} < o^{K,1*}$ for $K=G,B$, then this inequality does not translate into a difference concerning assignments in equilibrium. This proves iv).

We now consider the assignment of young workers to jobs. y^{K*} is the value for y such that a young worker is equally productive in the two jobs. That is, y^{K*} satisfies (A11).

$$(A11) \quad d_1+c_1[y^{K*}f(0)]=d_2^K+c_2[y^{K*}f(0)]$$

Let $E(W_O^J | S)$ be the expected productivity of an old worker of schooling level S who was assigned to job J when she was young (where this expectation is taken prior to the realization of the state of the world when the worker is old). Because being assigned to job 1 when young only counts of a period in the human-capital-accumulation function if the worker is assigned to job 2 when old, we know $E(W_O^2 | S) > E(W_O^1 | S)$. Given this, $y^K(S)$ satisfies (A12).

$$(A12) \quad d_1+c_1[y^K(S)f(0)]+E(W_O^1 | S)=d_2^K+c_2[y^K(S)f(0)]+E(W_O^2 | S)$$

(A11) yields $y^{K*}=(d_1-d_2^K)/(c_2-c_1)f(0)$ while (A12) yields $y^K(S)=[d_1-d_2^K+E(W_O^1 | S)-E(W_O^2 | S)]/(c_2-c_1)f(0)$. A comparison of these expressions yields $y^K(S) < y^{K*}$ for all S and $K=G,B$. This proves i).

Further, the expression for $y^K(S)$ also yields $y^G(S) < y^B(S)$ for all S , while the earlier expressions for $o^{K,2*}$ and $o^{K,1*}$ yield $o^{G,J}(S) < o^{B,J}(S)$ for all S and $J=1,2$. This proves iii). Q.E.D.

Proof of Proposition 3: Competition and no firm-specific human capital means that a worker's wage each period equals her expected output in the job to which she is assigned. Let Z_n denote the number of workers of schooling level n in each cohort, and let Z denote the total number of workers in each cohort. Given a worker is always paid her expected output in the job to which she is assigned, we have that W_Y^B must satisfy (A13).

$$(A13) \quad W_Y^B < Z_n \max\{d_1+c_1E(y | S_n)f(0), d_2^B+c_2E(y | S_n)f(0)\}/Z$$

Now consider the good state of the world. y^{G*} satisfies $d_1+c_1 y^{G*}f(0)=d_2^G+c_2 y^{G*}f(0)$. Given $c_2 > c_1$, sufficiently high values for d_2^G yield $y^{G*} < L+B(S)$ for all S . But since Proposition 2 tells us that $y^G(S) < y^{G*}$ for all S , we now have that $y^G(S) < L+B(S)$ for all S for sufficiently high values for d_2^G . For such a value for d_2^G we can write W_Y^G as in (A14).

$$(A14) \quad W_Y^G = Z_n(d_2^G+c_2E(y | S_n)f(0))/Z$$

Comparing (A13) and (A14), in turn, yields that $W_Y^G > W_Y^B$ for sufficiently high values for d_2^G .

We now turn our attention to the average wage of old workers in state of the world K . Consider first a worker of schooling level S such that $E(y | S) = y^B(S)$. Given from Proposition 2 we know

$y^G(S) < y^B(S)$ for all S , when she is young this worker is assigned to the high-level job in both states of the world. In turn, since job assignments and wages when a worker is old depends on the current state of the world, the previous job assignment, and the current expectation concerning the worker's ability to learn on the job (and not directly on the previous period's state of the world), the worker's expected wage when she is old is independent of whether the state of the world when the worker was young was good or bad. Further, a similar argument yields that the expected wage when a worker is old for a worker of schooling level S such that $E(\epsilon_i | S) < y^G(S)$ is independent of whether the state of the world when the worker was young was good or bad.

Now consider a worker of schooling level S such that $y^G(S) > E(\epsilon_i | S) > y^B(S)$. When this worker is young she will be assigned to the high-level job in the good state of the world and the low-level job in the bad state. In turn, if the state of the world when she was young was good rather than bad, then her expected wage when she is old is higher. The logic here is as follows. First, given $\beta < 1$, if she is assigned to the high-level job in both states of the world she earns more when the previous period's state of the world was good. Second, given that there is no task-specific human capital on job 1 (see equation (14)), if she is assigned to the low-level job in both states of the world her pay when old is independent of the previous period's state of the world. Third, since $o^{K,2}(S) < o^{K,1}(S)$, she is sometimes assigned to the high-level job when the previous period was good but assigned to the low-level job when the previous period was bad. Further, whenever this is the case, since $o^{K,2}(S)$ satisfies $d_1 + c_1 o^{K,2}(S) f(1) = d_2^K + c_2 o^{K,2}(S) f(1)$ and $c_2 > c_1$, she is paid more when the previous period's state of the world was good. In summary, her expected wage when old is higher when the previous period's state of the world was good since she always earns at least as much when the previous period was good rather than bad and she sometimes earns more.

We now have that a worker's expected wage when she is old is independent of the state of the world in the previous period when $E(\epsilon_i | S_i) > y^B(S_i)$ and $E(\epsilon_i | S_i) < y^G(S_i)$, while when $y^G(S_i) > E(\epsilon_i | S_i) > y^B(S_i)$ the worker's expected wage when she is old is higher when the state of the world in the previous period was good. Given that there is at least one schooling group such that $E(\epsilon_i | S_i)$ falls in the latter interval (see footnote 24), we now have $W_O^{K,G} > W_O^{K,B}$ for $K=G,B$. This proves ii). Q.E.D.

Proof of Proposition 4: Using the same logic as in the proofs of Propositions 2 and 3 it must be the case that $y^G(S) < y^B(S)$ for all S and that there exists a value d_2^G such that, if $d_2^G > d_2^B$, then the average wage for young workers is higher when the state of the world is good rather than bad. Now consider the assignment of workers to jobs and wages over the last $T-1$ periods a cohort is in the labor market. Consider first a worker of schooling level S such that $E(\epsilon_i | S) > y^B(S)$. Given $y^G(S) < y^B(S)$ for all S , in her first period in the labor market this worker is assigned to the high-level job in both states of the world. In turn, since the state of the world when a worker is young does not directly enter into the production function for either job in later periods, we have that the worker's expected compensation in

the next $T-1$ periods must be independent of the state of the world when she was young. Further, a similar argument yields that, for a worker of schooling level S such that $E(\epsilon_i | S) < \gamma^G(S)$, the worker's expected compensation over her last $T-1$ periods in the labor market is independent of the state of the world when she was young.

Now consider a worker of schooling level S such that $\gamma^G(S) < E(\epsilon_i | S) < \gamma^B(S)$. When this worker is young she will be assigned to the high-level job in the good state of the world and the low-level job in the bad state. In turn, if the state of the world when she was young was good rather than bad, then the worker's expected compensation over her last $T-1$ periods in the labor market must be higher because of the following. Suppose the state of the world when the worker was young was good. In this case, for every sequence of realizations of the stochastic term, the worker has the option in each period of choosing the same job assignment that she chooses when the state of the world when she was young was bad rather than good. Since, in turn, some of these job assignments are to the high-level job (see footnote 26) and because $\beta < 1$, this set of choices must result in a higher value for the worker's expected compensation over her last $T-1$ periods in the labor market than she achieves when the state of the world was bad when she was young. But since her actual choices must do at least as well as this, we now have that the worker's expected compensation over her last $T-1$ periods in the labor market must be higher when the state of the world when she was young was good rather than bad.

We now have that a worker's expected compensation over her last $T-1$ periods in the labor market is independent of the state of the world when the worker was young when $E(\epsilon_i | S_i) = \gamma^B(S_i)$ and $E(\epsilon_i | S_i) < \gamma^G(S_i)$, while when $\gamma^G(S_i) < E(\epsilon_i | S_i) < \gamma^B(S_i)$ this expected compensation is higher when the state of the world when the worker was young was good rather than bad. Given that there is at least one schooling group such that $E(\epsilon_i | S_i)$ falls in the latter interval (see footnote 24), we now have that a cohort's average compensation over the last $T-1$ periods it is in the labor market is higher when the state of the world was good rather than bad in its first period in the labor market. This proves i).

Suppose $\beta < 1$ and consider a worker who has spent her first $T-1$ periods in the labor market in job 1. Let ϵ be her expected ability to learn on the job. Since $\beta < 1$, we know $d_1 + c_1 \beta^{T-1} > d_2 + c_2 \beta^{T-1}$. Thus, she will be assigned to the low-level job in her last period in the labor market. In turn, this immediately yields that there must exist a smallest value t , $t \leq T-1$, such that a worker who spends her first t periods in job 1 is not promoted in any of her last $T-t$ periods in the labor market. This proves ii).

We demonstrate the third point by presenting an example in which some promotions are associated with wage decreases. Suppose $T=3$, $d_1=20$, $c_1=\beta$, $d_2=10$, $c_2=1$, $\beta=1/2$, $f(0)=\beta$, $f(1/2)=9$, $f(3/2)=16$, where β is infinitesimally small. Also, we consider a schooling group for which $\beta_L + B(S) = \beta$, and $\beta_H + B(S) = 1$ which, given $d_1=20$, $d_2=10$, and $f(0)=\beta$, yields that a worker in this schooling group is assigned to the low-level job in her first period in the labor market. Let W_t denote the worker's wage in her t^{th} period in the labor market. We have $W_1 = 20$ since $c_1 = \beta$ and $d_1 = 20$. Now suppose that learning after

the first period is such that the expectation in the second period concerning the worker's ability to learn on the job is arbitrarily close to $\frac{1}{2}H+B(S)$. If the worker remains in job 1 for the remaining two periods then she earns $W_2+W_3 = 20+20=40$ over these last two periods since $c_1=$ and $d_1=20$. If instead she is promoted, then $W_2+W_3 = 19+26=45$ since her expected ability to learn on the job is arbitrarily close to $\frac{1}{2}H+B(S)$. Hence, she will be promoted at the beginning of the second period but her wage falls from approximately 20 to approximately 19 when the promotion takes place. This proves iii). Q.E.D.

REFERENCES

- Altonji, J. and C. Pierret (2001), "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116, pp. 313-350.
- Baker, G., M. Gibbs, and B. Holmstrom (1994a), "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 109, pp. 881-919.
- Baker, G., M. Gibbs, and B. Holmstrom (1994b), "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109, pp. 921-955.
- Baker, M. (1997), "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15, pp. 338-375.
- Beaudry, P. and J. DiNardo (1991), "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Microdata," *Journal of Political Economy*, 99, pp. 665-688.
- Becker, G. (1962), "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70, pp. 9-49.
- Becker, G. (1964), *Human Capital* (NBER, New York).
- Bernhardt, D. (1995), "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, pp. 315-339.
- Burdett, K. (1978), "A Theory of Employee Search and Quit Rates," *American Economic Review*, 68, pp. 212-220.
- Card, D. (1999), "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 (North-Holland, Amsterdam), pp. 1801-1863.
- Chaudhury, P. (2002), "Multi-tasking and the Returns to Experience," mimeo, New York University.
- Devereux, P. (2002a), "State Dependence in Wages: Stigma, Human Capital, or Search," mimeo, UCLA.
- Devereux, P. (2002b), "Occupational Upgrading and the Business Cycle," *Labour*, 16, pp. 423-452.
- Doeringer, P. and M. Piore (1971), *Internal Labor Markets and Manpower Analysis* (Heath Lexington Books, Lexington, MA).
- Farber, H. and R. Gibbons (1996), "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111, pp. 1007-1047.
- Ghosh, S. (2002), "Job Mobility and Careers in Firms," mimeo, ITAM Business School.
- Gibbons, R. (1997), "Incentives and Careers in Organizations," in D. Kreps and K. Wallis, eds., *Advances in Economic Theory and Econometrics* (Cambridge University Press, New York), pp. 1-37.
- Gibbons, R. and L. Katz (1991), "Layoffs and Lemons," *Journal of Labor Economics*, 9, pp. 351-380.
- Gibbons, R. and M. Waldman (1999a), "A Theory of Wage and Promotion Dynamics Inside Firms," *Quarterly Journal of Economics*, 114, pp. 1321-1358.

- Gibbons, R. and M. Waldman (1999b), "Careers in Organizations: Theory and Evidence," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 (North-Holland, Amsterdam), pp. 2373-2437.
- Greenwald, B. (1986), "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53, pp. 325-347.
- Habermalz, S. (2002), "Job Matching and the Returns to Education Signals," IZA Discussion Paper No. 726, The Institute for the Study of Labor (IZA) in Bonn.
- Harris, M. and B. Holmstrom (1982), "A Theory of Wage Dynamics," *Review of Economic Studies*, 49, pp. 315-333.
- Heckman, J., L. Lochner, and R. Cossa (2002), "Learning-By-Doing Vs. On-the-job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation," NBER Working Paper No. 9083.
- Jovanovic, B. (1979), "Job Mobility and the Theory of Turnover," *Journal of Political Economy*, 87, pp. 972-990.
- Jovanovic, B. and Y. Nyarko (1997), "Stepping Stone Mobility," Carnegie-Rochester Series for Public Policy, 46, pp. 289-337.
- Kahn, L. (2002), "The Labor Market Consequences of Graduating College in a Bad Economy," mimeo, University of Chicago.
- Kwon, I. (1998), "Incentives, Wages, and Promotions: Theory and Evidence," mimeo, Harvard University.
- Lazear, E. (1992), "The Job as a Concept," in W. Bruns, ed., *Research in Labor Economics*, Vol. 8 (JAI Press, Greenwich, CT), pp. 141-165.
- Lazear, E. (2002), "The Peter Principle: Promotions and Declining Productivity," mimeo, Stanford University (Forthcoming *Journal of Political Economy*).
- Lazear, E. and S. Rosen (1981), "Rank-Order Tournaments and Optimum Labor Contracts," *Journal of Political Economy*, 89, pp. 841-864.
- Lima, F. and P. Pereira (2001), "Career Progression within Large Firms," mimeo, Faculdade de Economia da Universidade Nova de Lisboa.
- Lluis, S. (2001), "The Role of Comparative Advantage and Learning in Wage Dynamics and Intrafirm Mobility: Evidence from Germany," mimeo, University of Montreal.
- MacLeod, B. and J. Malcomson (1993), "Investments, Holdup, and the Form of Market Contracts," *American Economic Review*, 83, pp. 811-837.
- Main, B., C. O'Reilly III, and J. Wade (1993), "Top Executive Pay: Tournaments or Teamwork?" *Journal of Labor Economics*, 11, pp. 606-628.
- McCue, K. (1996), "Promotions and Wage Growth," *Journal of Labor Economics*, 14, pp. 175-209.
- Medoff, J. and K. Abraham (1980), "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95, pp. 703-736.

- Medoff, J. and K. Abraham (1981), "Are Those Paid More Really More Productive," *Journal of Human Resources*, 16, pp. 186-216.
- Mincer, J. (1974), *Schooling, Experience, and Earnings* (NBER, New York).
- Mortensen, D. (1978), "Specific Capital and Labor Turnover," *Bell Journal of Economics*, 9, pp. 572-586.
- Munasinghe, L. (2001), "A Theory of Wage and Turnover Dynamics," mimeo, Columbia University.
- Murphy, K. (1985), "Corporate Performance and Managerial Remuneration: An Empirical Analysis," *Journal of Accounting and Economics*, 7, pp. 11-42.
- Neal, D. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13, pp. 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, pp. 306-323.
- Podolny, J. and J. Baron (1997), "Resources and Relationships: Social Networks and Mobility in the Workplace," *American Sociological Review*, 62, pp. 673-693.
- Prendergast, C. (1993), "The Role of Promotion in Inducing Human Capital Acquisition," *Quarterly Journal of Economics*, 108, pp. 523-534.
- Ricart i Costa, J. (1988), "Managerial Task Assignments and Promotion," *Econometrica*, 56, pp. 449-466.
- Rosen, S. (1982), "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13, pp. 311-323.
- Rosen, S. (1986), "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76, pp. 701-715.
- Rosenbaum, J. (1984), *Career Mobility in a Corporate Hierarchy* (Academic Press, New York).
- Topel, R. and R. Ward (1992), "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107, pp. 439-480.
- Waldman, M. (1984a), "Worker Allocation, Hierarchies, and the Wage Distribution," *Review of Economic Studies*, 51, pp. 95-109.
- Waldman, M. (1984b), "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, 15, pp. 255-267.
- Waldman, M. (1990), "Up-Or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8, pp. 230-250.
- Weiss, Y. (1991), "Learning by Doing and Occupation Specialization," *Journal of Economic Theory*, 3, pp. 189-198.
- Welch, F. (1979), "Effects of Cohort Size on Earnings: The Baby Boom Baby's Financial Bust," *Journal of Political Economy*, 87, pp. S65-S97.

Willis, R. (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, Vol. 1 (North-Holland, Amsterdam), pp. 525-602.

Zabojnik, J. and D. Bernhardt (2001), "Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation," *Review of Economic Studies*, 68, pp. 693-716.