A THEORY OF WAGE AND PROMOTION DYNAMICS INSIDE FIRMS*

ROBERT GIBBONS AND MICHAEL WALDMAN

We show that a framework that integrates job assignment, human-capital acquisition, and learning captures several empirical findings concerning wage and promotion dynamics inside firms, including the following. First, real-wage decreases are not rare but demotions are. Second, wage increases are serially correlated. Third, promotions are associated with large wage increases. Fourth, wage increases at promotion are small relative to the difference between average wages across levels of the job ladder. Fifth, workers who receive large wage increases early in their stay at one level of the job ladder are promoted quickly to the next.

I. INTRODUCTION

A recurrent theme in labor economics is that careers in organizations deviate from the predictions of the standard theory of competitive labor markets. Doeringer and Piore [1971], for example, argue that wage rates are often more closely associated with job assignments than with workers' human-capital characteristics. More recently, Medoff and Abraham [1980, 1981] and Baker, Jensen, and Murphy [1988] have also presented evidence on a variety of practices that seem at odds with the standard competitive model. Continuing in this vein, recent papers by Lazear [1992] and Baker, Gibbs, and Holmstrom [1994a, 1994b] provide detailed empirical analyses of careers in particular firms.

In response to evidence presented by these and other authors (discussed in more detail below), a literature has developed that attempts to provide a theoretical foundation for several common features of careers in organizations. We focus on three principal approaches from this literature: job assignment, on-the-job human-capital acquisition, and learning. Our paper integrates these

* We thank George Baker, Michael Gibbs, Bengt Holmstrom, Lawrence Katz, Caroline Minter Hoxby, two anonymous referees, and seminar audiences at Chinese University of Hong Kong, Harvard University, Hong Kong University of Science and Technology, Massachusetts Institute of Technology, University of California Santa Cruz, University of Chicago, University of Toronto, and Society of Labor Economics for helpful comments. For financial support Gibbons thanks the Massachusetts Institute of Technology's Sloan School, Cornell University's Johnson School, the NSF (SBR-9809107), and the Center for Advanced Study in the Behavioral Sciences (for a fellowship funded in part by NSF grant SBR-9022192), while Waldman thanks Cornell University's Johnson School and the George J. Stigler Center for the Study of the Economy and the State at the University of Chicago.

The Quarterly Journal of Economics, November 1999

three theoretical approaches in a natural fashion, in an attempt to explain the main findings of the recent empirical literature. We begin by giving brief descriptions of these three theoretical approaches; see Gibbons and Waldman [1999] for a survey.

By the job-assignment literature we mean papers that investigate the assignment of workers to jobs when firms consist of a variety of potential job assignments and there is full information about workers and jobs. The models in this literature are typically static and focus on the roles of comparative advantage and the "scale-of-operations" effect (i.e., the idea that workers with higher ability should be assigned to jobs where decisions have an impact over a larger scale of operations) in determining the equilibrium assignment. For example, Sattinger [1975] considers a model where comparative advantage is the determining factor, while Rosen [1982] and Waldman [1984a] focus on the role of the scale-of-operations effect. This literature has produced a number of insights concerning the wage distribution both within and across firms; see Sattinger [1993]. Because the models are static, however, they do not provide direct explanations for wage and promotion dynamics.

A second important perspective on careers in organizations involves on-the-job human-capital acquisition. Since Becker [1964], an extensive literature has developed the implications of on-the-job human-capital acquisition for age-earnings profiles. Examples include Ben-Porath's [1967] analysis of how the incentive to invest in human capital varies over a worker's career and Hashimoto's [1981] analysis of the financing decision. More recently, a number of authors have combined the human-capital perspective with promotion and (in some cases) job assignment. For example, both Carmichael [1983] and Prendergast [1993] show how promotion can be used to provide incentives for efficient human-capital acquisition, while Kahn and Huberman [1988] explore the role of Up-or-Out contracts in providing such incentives.

A recent third literature investigates the role of learning. In this approach, firms are uncertain about a worker's ability when the worker enters the labor force, but gradually learn about the worker's ability during his or her career. Such papers typically fall into one of two categories. One set of papers assumes symmetric learning; that is, any information generated about a worker's ability during the worker's career is public information. Examples include Harris and Holmstrom's [1982] analysis of insurance, Holmstrom's [1982] model of career concerns, and Farber and

Gibbons' [1996] investigation of wage dynamics. The other category is asymmetric learning. These papers assume that a worker's current employer receives better information about the worker's ability than do prospective employers. Examples include Greenwald's [1986] model of adverse selection in the labor market, the analyses of Waldman [1984b, 1990], Ricart i Costa [1988], and Bernhardt and Scoones [1993], in which promotions serve as a signal of ability, and Gibbons and Katz's [1991] application of the adverse-selection and signaling approaches to data on the consequences of layoffs versus plant closings.

In this paper we develop a model that integrates job assignment, on-the-job human-capital acquisition, and learning.[1] We show that a framework that integrates these familiar ideas captures a number of recent empirical findings concerning wage and promotion dynamics inside firms, including but not limited to many of those found in Baker, Gibbs, and Holmstrom (hereinafter, BGH). In their study, BGH consider the ability of a variety of theoretical approaches to explain their empirical findings. Considering the human-capital-acquisition and learning approaches separately, they conclude that their empirical findings allow them to reject both approaches as explanations for wage and promotion practices in their firm. Our conclusion, in contrast, is that a model that combines a number of existing approaches explains many of BGH's findings. Hence, rather than ruling out existing theoretical approaches to wage and promotion practices inside firms, our interpretation is that the BGH (and other) findings support many of these approaches—although in a more integrated form than typically appears in the literature.

This paper contributes to a growing literature that attempts to provide a theoretical explanation for a broad pattern of evidence, rather than a model focused on a single empirical finding. Previous papers in this "broad pattern" spirit include Harris and Holmstrom's [1982] analysis of insurance, MacLeod and Malcomson's [1988] dynamic model of adverse selection and moral hazard, and Demougin and Siow's [1994] study of on-the-job training and screening. The paper closest to ours is Bernhardt

1. A number of the papers referred to above combine two or more of these elements. For example, Prendergast [1993] considers both job assignment and human-capital acquisition, while Waldman [1984b] and Bernhardt and Scoones [1993] capture aspects of all three elements. Other papers that combine two or more of the modeling elements we consider include Bernhardt [1995] and Jovanovic and Nyarko [1997]. With the exception of Bernhardt [1995], none of these papers addresses broad patterns of evidence.

[1995], which also considers a model that combines job assignment, on-the-job human-capital acquisition, and learning. As we discuss in Section V, we focus on symmetric learning while Bernhardt considers asymmetric learning, and there are several differences in the patterns of evidence that the resulting models can explain.

The outline for the paper is as follows. Section II reviews the recent empirical evidence on wage and promotion dynamics. Section III combines job assignment and on-the-job human-capital acquisition in a model with a three-level job ladder and full information. Section IV analyzes this model under the assumption of symmetric learning. Section V relates our framework to Medoff and Abraham's findings concerning performance evaluations, discusses findings from the BGH study that are inconsistent with the predictions of our symmetric-learning model, and discusses other models of wage dynamics inside firms. Section VI concludes.

II. Evidence on Wage and Promotion Dynamics

The main focus of Sections III and IV is the extent to which our theoretical framework explains the findings of BGH. (Subsection V.B discusses the major BGH findings that we do not capture.) We proceed in this fashion not because BGH is the only empirical study that considers wage and promotion dynamics inside firms, but rather because various studies have investigated different issues in different environments, so it is difficult to assess what combination of facts is true in any given environment. We therefore focus on a single study, choosing BGH because it is the most comprehensive. But the fact that our framework can explain many of the BGH findings is of more interest if these findings are representative of other authors' findings concerning wage and promotion dynamics. In this section, therefore, we discuss the extent to which the BGH findings that we focus on in Sections III and IV are representative of the larger literature. See Gibbons [1997] for a more comprehensive review of this literature.

BGH analyze personnel data from a single firm over a twenty-year time period. An observation in their data set consists of a managerial employee in a specific year. The information they have for that manager includes the employee's ID number, job title, salary, and performance rating. The first step of their analysis was to use the movement of managers across job titles to define a single job ladder consisting of eight distinct levels. Throughout our discussion of their evidence we mean by a promotion or a demotion a movement up or down this job ladder. It would be interesting to know how many of these promotions and demotions involve changes only in job title, not in job content, but the data shed no light on this.

The BGH findings we concentrate on are the following. First, real-wage decreases are not rare, while demotions are very rare. Second, there is significant serial correlation in both wage increases and promotion rates (the latter is sometimes referred to as a "fast track"). Third, promotions are associated with large wage increases. Fourth, wage increases at promotion are small relative to the difference between average wages across levels of the job ladder. Fifth, workers who receive larger wage increases early in their stay at one level of the job ladder are promoted more quickly to the next level.

To appreciate the context of the BGH study, it is useful to consider the range of other studies that address one or more of these findings. Some of these studies, such as Rosenbaum [1984] and Lazear [1992], are similar to BGH in that they look at careers inside particular firms, while other studies such as Abowd and Card [1989], Topel [1991], and Baker [1997] analyze data that include mobility between firms. Another way that the studies differ is that some (like BGH) focus on managerial and professional workers, while others analyze broader cross sections of workers. Our goal is to develop a model consistent with the evidence from BGH concerning managers from a single firm. Our model may also apply in other settings, but there are clearly some settings in which our model will not apply. For example, the rules regarding promotions in a unionized establishment are very different from the predictions of our model.

Two findings with strong support in the literature are that real-wage decreases are not rare and that there is serial correlation in promotion rates. For example, McLaughlin [1994] and Card and Hyslop [1997] both find a significant frequency of real-wage decreases, while there is extensive support for serial correlation in promotion rates (e.g., Rosenbaum [1984], Bruderl, Diekmann, and Preisendorfer [1991], and Podolny and Baron [1997]). Further, although we know of no study besides BGH that documents this finding, we take it to be uncontroversial that demotions are quite rare.

The evidence concerning serial correlation in wage increases is a bit less clear. Lillard and Weiss [1979], Hause [1980], and Baker [1997] find evidence of serially correlated wage changes (or wage residual changes), while Abowd and Card [1989], Topel [1991], and Topel and Ward [1992] do not. One possible explanation relies on differences in the samples studied. BGH, Lillard and Weiss, and Hause all study relatively homogeneous samples (BGH managers in a single firm, Lillard and Weiss American scientists, and Hause young Swedish males), while the other studies investigate large heterogeneous samples such as the Panel Study of Income Dynamics (PSID). Another possible explanation involves panel length and the power of alternative econometric tests. Baker finds serial correlation in a twenty-year panel from the PSID and argues that earlier studies based on similar samples did not find serial correlation because of either short panel length or low-powered tests.

Another finding with strong support in the literature is that promotions are associated with large wage increases; see Gerhart and Milkovich [1989], Lazear [1992], and McCue [1996], among others. There is also evidence that these wage increases, although large relative to wage increases associated with not being promoted, are small relative to the difference between average wages across levels of a job ladder. Murphy [1985] documents both of these facts: in a study of top executives in 72 large U. S. manufacturing firms, he finds that the average real increase in salary plus bonus for the whole sample was 3.7 percent, the average increase for a vice president promoted to president was 21 percent, and the average salary plus bonus for presidents was 60 percent higher than for vice presidents. See also Main, O'Reilly, and Wade [1993].

In sum, the first four BGH findings we focus on find substantial support in the literature. In contrast, there is little in the literature concerning the remaining finding—that workers who receive larger wage increases early in their stay at one level of a job ladder are promoted more quickly to the next level. As far as we know, no other study directly addresses this issue. However, McCue [1996] does find something similar: a high wage today is positively correlated with promotion tomorrow, which suggests that the BGH finding might hold in McCue's data set.

III. An Analysis of Job Ladders under Full Information

This section develops a model in which firms consist of three-level job ladders. Job-assignment and human-capital considerations determine wage and promotion dynamics. As a benchmark, in this section we analyze this model under full information; the following section provides an analysis under symmetric learning. The technology of production we consider throughout the paper is similar to that investigated in Sattinger [1975], Rosen [1982], and Waldman [1984a]: jobs are ranked in terms of the extra value produced by a worker of greater ability. Sattinger, Rosen, and Waldman consider this technology in one-period models characterized by full information and no human-capital acquisition, while our focus is on multiple periods, human-capital acquisition, and (in later sections) learning.

A. The Model

There is free entry into production. All firms are identical, and the only input is labor. A worker's career lasts for T periods, $T \geq 5$ (an unusual assumption which we motivate below). In each period, labor supply is fixed at one unit for each worker. Worker i's *innate ability* is denoted θ_i and can be either high or low: $\theta_i \in \{\theta_H, \theta_L\}$. A worker's *effective ability* is a function of the worker's innate ability and the worker's labor-market experience. Let t denote calendar time and x_{it} the worker's labor-market experience prior to period t (i.e., for a worker in his or her first period in the labor market, prior experience x_{it} equals zero). We assume that worker i's effective ability in period t is given by

$$\eta_{it} = \theta_i f(x_{it}), \tag{1}$$

where $f' > 0$ and $f'' \leq 0$.

A firm consists of three different jobs, denoted 1, 2, and 3. If worker i is assigned to job j in period t, then the worker produces

$$y_{ijt} = d_j + c_j(\eta_{it} + \epsilon_{ijt}), \tag{2}$$

where d_j and c_j are constants known to all labor-market participants and ϵ_{ijt} is a noise term drawn from a normal distribution with mean 0 and variance σ^2. Let η' denote the effective ability level at which a worker is equally productive at jobs 1 and 2. That is, η' solves $d_1 + c_1\eta = d_2 + c_2\eta$. Similarly, let η'' solve $d_2 + c_2\eta = d_3 + c_3\eta$. We assume that $c_3 > c_2 > c_1 > 0$ and $0 < d_3 < d_2 < d_1$, and that these parameters are such that $\eta'' > \eta'$. Thus, given full information about worker abilities, the efficient assignment rule for period t is to assign worker i to job 1 if $\eta_{it} < \eta'$, to job 2 if $\eta' < \eta_{it} < \eta''$, and to job 3 if $\eta_{it} > \eta''$.

This specification of the production technology parallels those in Waldman [1984b], Gibbons and Katz [1992], and Bernhardt

[1995], in that a worker's productivity in each job is independent of the assignment of other workers within the firm. Thus, taking the model literally, the firm's productivity might be optimized by assigning all the workers to one level of the job ladder. Rosen [1982], Waldman [1984a], and MacDonald and Markusen [1985] incorporate factors such as slot constraints and so can address issues such as congestion that we omit in order to keep our analysis tractable. We interpret equation (2) as a plausible reduced form for the productivity of an individual worker in a large firm in which many workers are assigned to each job level in the steady state.

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.

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 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, in the full-information case considered in this section and the symmetric-learning analysis of Section IV, there is no turnover in equilibrium.[2]

To reduce the number of cases that need to be considered, we restrict the analysis to parameterizations that satisfy the follow-

2. As an alternative to this infinitesimal moving cost borne by the worker, we could instead assume a hiring cost borne by the firm. That is, in the first period of a worker's employment at a firm, the worker's productivity is lower than specified in equation (2) by an amount h. If h is infinitesimal, then the results are equivalent to those we present. If h is significant, then all the results are the same except for the wage in the worker's final period. In this period prospective employers are willing to pay productivity as defined by equation (2) minus the hiring cost, so this is the worker's wage. In the next-to-last period, prospective employers are willing to pay the worker's productivity, as defined by equation (2), so now this is the worker's wage. The reason that firms are willing to pay this wage in the next-to-last period is that if such a firm succeeds in recruiting the worker, the firm will bear the cost h this period but earn a profit of h in the final period. This logic applies to all earlier periods, including the worker's first period in the labor market. If there were a discount rate r per period, the wage in the final period would be the same as above, but each earlier wage would be decreased by an amount $rh/(1 + r)$.

ing conditions. First, $\theta_H f(1) < \eta'$, so that it is efficient for each worker to be assigned to job 1 in the first two periods of the worker's career. (Recall that x_{it} measures *prior* labor-market experience, so $x = 1$ in the worker's second period in the labor market.) Second, $\eta'' - \eta' > \theta_H[f(3) - f(1)]$, so that it is efficient for each worker to be at job 2 for at least two periods before being promoted to job 3 (since $\theta_H > \theta_L$ and $f(3) - f(1) \geq f(x + 2) - f(x)$ for all $x \geq 1$). Third, $\theta_L f(T - 1) > \eta''$, so that it is efficient for each worker to be on job 3 by the end of the worker's career. For these three conditions to hold simultaneously requires that $T \geq 5$.

Finally, in the analyses that follow we focus on absolute rather than percentage wage changes although the empirical literature (including BGH) typically looks at percentage wage changes. Given our assumption that $d_3 > 0$, all but one of our results that concern wage changes hold for both absolute and percentage wage changes (and even in the remaining case the result holds for percentage wage changes for many parameterizations; see footnote 7).

B. The Full-Information Benchmark

We now analyze the benchmark case of full information. That is, each worker's innate ability (θ_i) is common knowledge at the beginning of the worker's career. If innate ability is common knowledge, then in each period a worker is assigned to the job that maximizes the worker's expected output and paid a wage equal to that expected output.[3] Proposition 1 characterizes the assignment rule in this case. We use w_{it} to denote the wage paid to worker i in period t in equilibrium. Proofs are given in the Appendix.

PROPOSITION 1. Suppose that each worker's innate ability is common knowledge at the beginning of the worker's career. Then job assignments and wages are given by (i) through (iii):
(i) If $\eta_{it} < \eta'$, then worker i is assigned to job 1 in period t and earns the wage $w_{it} = d_1 + c_1\eta_{it}$.
(ii) If $\eta' \leq \eta_{it} < \eta''$, then worker i is assigned to job 2 in period t and earns the wage $w_{it} = d_2 + c_2\eta_{it}$.
(iii) If $\eta_{it} \geq \eta''$, then worker i is assigned to job 3 in period t and earns the wage $w_{it} = d_3 + c_3\eta_{it}$.

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

Proposition 1 says that in equilibrium there is a job ladder that workers climb as they gain labor-market experience. That is, the equilibrium job assignments described in Proposition 1 can be restated in terms of the worker's labor-market experience rather than the worker's effective ability: for each θ_i ($i = H,L$) there exist values x'_i and x''_i, where $x''_i > x'_i$, such that if a worker's innate ability is θ_i then in period t the worker is assigned to job 1 if $x_{it} < x'_i$, to job 2 if $x'_i \leq x_{it} < x''_i$, and to job 3 if $x_{it} \geq x''_i$. There are no demotions in equilibrium because effective ability increases monotonically as a worker gains labor-market experience, so it is never optimal to promote and then demote a worker. For the same reason, a worker's wage rises every period.

The fact that there are no demotions in equilibrium is roughly consistent with the results of BGH, who found that 0.3 percent of the year-to-year job movements in their sample were demotions. In contrast, BGH found that approximately 25 percent of the year-to-year salary changes in their sample were real-wage decreases. Certain years have pronounced effects, however: in each of the high-inflation years of 1979 and 1980, the median real-salary increase was negative at the firm BGH investigated. But even in a typical year, when the average real-salary increase was between 5 percent and 9 percent, the fraction of year-to-year salary changes that were real-wage decreases never fell below 5 percent, and across the nineteen years of their study, the median value for this fraction was 12 percent.

BGH's evidence on demotions and (especially) real-wage decreases does not match the results of our full-information analysis. In contrast, the full-information case does quite well at capturing two of the other findings discussed in Section II: serial correlation in wage increases and promotion rates, and wage increases predict promotion. More specifically, BGH find serial correlation in wage increases after controlling for age, education, tenure, and salary. They also find that, holding constant tenure in level 2, promotion probability is higher for those promoted more quickly from level 1 (and similar patterns for higher levels). Finally, they find that those with larger raises in their first year in level 1 are promoted more quickly to level 2 (and similarly for initial raises in level 2 predicting promotion to level 3).

To understand the source of serial correlation of wage increases in our model, consider workers with labor-market experience x. The wage increase for a θ_i worker whose experience increases from x to $x + 1$ and who is not promoted is

$c_j\theta_i[f(x+1) - f(x)]$, where j denotes the worker's job level. This expression is larger for θ_H than θ_L because, given experience x, the θ_H worker is assigned to at least as high a job level j. For the same reasons, the wage increase for a θ_i worker whose experience increases from $x+1$ to $x+2$ and who is not promoted is again larger for θ_H than θ_L. Thus, wage changes in the absence of promotions are serially correlated after controlling for experience. A slightly more complicated argument shows that this result also holds when promotions are allowed.[4]

The model also captures serial correlation in promotion rates in that, if η' and $\eta'' - \eta'$ are both sufficiently large, then high-ability workers are promoted to job 2 more quickly and also spend less time on job 2 before being promoted to job 3. That is, $x'_H < x'_L$ and $x''_H - x'_H < x''_L - x'_L$.[5] The logic for this result is simple. A worker receives a promotion when effective ability reaches certain absolute levels. Since high-ability workers experience faster growth in effective ability, these workers are promoted to job 2 earlier in their career and are also promoted to job 3 after having spent less time on job 2.

Finally, to see why wage increases predict promotion, consider workers with labor-market experience x, where $x < x'_H$ so that the workers are in job 1, and assume that η' is sufficiently large that $x'_H < x'_L$. As above, holding experience fixed, high-ability workers receive larger wage increases than do low-ability workers. In turn, given that high-ability workers are also promoted to job 2 earlier in their careers, we have that wage increases predict promotion.

One of the other findings discussed in Section II concerns large wage increases upon promotion. This finding is captured by the full-information case in a weak form. To see why we say this, consider the average wage increase received by workers who are promoted to job 2. If f'' is close to zero (for all x), then the average wage increase received by workers promoted to job 2 is larger than the average wage increase received by workers who remain in job 1, because increases in effective ability are valued at rate c_1 in job

4. BGH also find serial correlation in wage increases without controlling for age, education, tenure, and salary. We conjecture that this will also hold in our model for some parameterizations. However, showing this would require a complicated analysis because whether it holds depends on several factors in addition to the factor given in the text, such as the concavity of the human-capital function $f(x)$ and the differences among the slopes c_1, c_2, and c_3.

5. If x were continuous, then $x'_H < x'_L$ and $x''_H - x'_H < x''_L - x'_L$ would hold for all parameterizations, rather than only if η' and $\eta'' - \eta'$ are sufficiently large.

1 but at rate $c_2 > c_1$ in job 2. The wage increase at promotion is the sum of two parts: c_1 times the worker's increase in effective ability, plus the increased value from assigning a worker with the new effective ability to job 2 rather than job 1. Therefore, the average wage increase at promotion to job 2 will be larger than the average wage increase received by workers who remain on job 1.

One reason we say the full-information case captures the empirical finding of a large wage increase upon promotion in a weak form is that we do not feel that the effect just described is by itself a plausible explanation for the findings of BGH. According to the argument just given, the average wage increase the year after a promotion to job 2 should be larger than the average wage increase at promotion to job 2 (because in the year after the promotion to job 2 increases in effective ability are valued at rate c_2 rather than at a convex combination of c_1 and c_2). But BGH find that for workers promoted from level 1 to level 2, the average wage increase the year after the promotion is strictly less than the average wage increase at promotion.[6]

The final finding that holds in the full-information case is that, although wage increases upon promotion are large (in a weak sense), they explain only a fraction of the difference between average wages across levels of a job ladder. Consider the difference between average wages at levels 1 and 2. Because workers at level 2 are on average older than workers at level 1, an important component of this difference is the increased productivity due to human-capital accumulation. Since the average wage increase at promotion will capture only one year of this age difference, the difference between average wages at levels 1 and 2 is bigger than the average wage increase at promotion to level 2. In fact, if the difference in average ages across levels is much greater than one, then the average wage increase at promotion can be very small relative to the difference between average wages across levels.

Overall, several results in this section are consistent with the BGH evidence, including serial correlation in both wage increases and promotion rates, and that workers who receive larger wage increases early in their stay at one level of the job ladder are promoted more quickly to the next. But the model does not capture a number of other findings discussed in Section II, including the nontrivial frequency of real wage decreases, the

6. BGH do not report results that allow us to make this comparison for promotion to any other level.

existence of demotions, and the size of wage increases upon promotion. In the next section we show that adding learning to the model improves its performance: the addition of learning leaves unchanged most of the predictions of the full-information analysis that match the BGH findings, but changes a number of the other predictions so that they match the evidence better.

IV. An Analysis of Job Ladders under Symmetric Learning

In this section we depart from the assumption of full information. In particular, when a worker enters the labor force, firms are now uncertain about the worker's ability and learn about it only gradually as the worker's career progresses. In this section we consider the case of symmetric learning: at any point in time all firms (and the worker in question) are equally informed about the worker's ability level. Our analysis in this section extends those of Murphy [1986] and Gibbons and Katz [1992], both of which study symmetric learning and assignment to a job ladder, but neither of which allows human-capital accumulation. (See also Ross, Taubman, and Wachter [1981] and MacDonald [1982].) Because our model also incorporates human-capital acquisition, we capture the idea that workers move up a job ladder as they age.

A. The Model and Preliminary Results

The only difference between this model and the full-information model in Section III is that now there is symmetric learning. At the beginning of a worker's career, the worker is known to be of innate ability θ_H with probability p_0 and of innate ability θ_L with probability $(1 - p_0)$. Learning takes place at the end of each period when the realization of the worker's output for that period becomes common knowledge. The presence of the noise term ϵ_{ijt} in the production function (2) implies that learning occurs gradually.

Define $z_{it} = (y_{ijt} - d_j)/c_j = \eta_{it} + \epsilon_{ijt}$. That is, z_{it} is the signal about the worker's effective ability 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 to $(z_{it-x}, \ldots, z_{it-1})$ as the worker's normalized output history at date t. Because the signal z_{it} is independent of job assignment, there is no difference in the rate of learning across jobs.

Let θ_{it}^e denote the expected innate ability of worker i in period t: $\theta_{it}^e = E(\theta_i | z_{it-x}, \ldots, z_{it-1})$. From θ_{it}^e we can compute the expected effective ability of worker i in period t:

$$\eta_{it}^e = \theta_{it}^e f(x_{it}). \tag{3}$$

Given η_{it}^e, job assignment and wage determination proceed as in the full-information case: in each period a worker is assigned to the job that maximizes the worker's expected output and is paid a wage equal to that expected output.

PROPOSITION 2. Suppose that learning is symmetric and that at the beginning of a worker's career the worker is known to be of innate ability θ_H with probability p_0 and of innate ability θ_L with probability $(1 - p_0)$. Then job assignments and wages are given by (i) through (iii):

(i) If $\eta_{it}^e < \eta'$, then worker i is assigned to job 1 in period t and $w_{it} = d_1 + c_1\eta_{it}^e$.

(ii) If $\eta' \leq \eta_{it}^e < \eta''$, then worker i is assigned to job 2 in period t and $w_{it} = d_2 + c_2\eta_{it}^e$.

(iii) If $\eta_{it}^e \geq \eta''$, then worker i is assigned to job 3 in period t and $w_{it} = d_3 + c_3\eta_{it}^e$.

Proposition 2 says that wages and job assignments are now determined by expected effective ability, whereas in the full-information case the actual value of effective ability determined wages and job assignments. Only the expected effective ability matters, rather than other moments of the distribution, because of the simple structure of our model. Specifically, output is a linear function of effective ability on each job and the rate of learning is independent of job assignment. We therefore cannot address certain situations such as where young workers are assigned to jobs that are unusually informative about their abilities.

B. Further Analysis

In this subsection we show that, despite the similarity between Propositions 1 and 2, the symmetric-learning case does a better job of matching the empirical evidence. We begin with the predictions of the full-information model that match the evidence well. In the full-information model, and in the data, there is serial correlation in both wage increases and promotion rates. As shown in Corollary 1, for workers in job 1 in periods t and $t + 1$, serial correlation continues to hold in the symmetric-learning case for wage changes at $t + 1$ and $t + 2$.[7]

COROLLARY 1. Consider a worker in period t who has x periods of labor-market experience and is currently in job 1 earning the wage w_{it}. If the worker is again in job 1 in period $t + 1$, then the conditional expectation of $w_{it+2} - w_{it+1}$ is an increasing function of $w_{it+1} - w_{it}$.

The logic for why there is serial correlation in wage increases in the situation considered in Corollary 1 is based on the logic in the full-information case. Starting at a given experience level and wage, a large wage increase indicates that an increase in expected innate ability occurred. But we know from the full-information case that wage increases are an increasing function of innate ability. The corollary follows because in the symmetric-learning case wage increases are an increasing function of expected innate ability.

Corollary 1 does not apply to workers who start on job 1 but whose first wage increase is associated with a promotion, or to workers who start on job 2 or job 3. We can generalize the result if we restrict the analysis to parameterizations for which there is no possibility of demotion.[8] That is, suppose that for any x, $\theta_H f(x) > \eta'$ implies $\theta_L f(x+1) > \eta'$ and $\theta_H f(x) > \eta''$ implies $\theta_L f(x+1) > \eta''$. Then for a worker in job j in period t, the expectation of $w_{it+2} -$

7. As discussed earlier, all but one of our results that concern wage changes hold for both absolute and percentage wage changes. The one result that we have been unable to prove for percentage wage changes is Corollary 1. In a proof available upon request, we show that the corollary holds for percentage wage changes given either $\theta_H f(x+2) \leq \eta'$ or $\theta_L f(x+2) \geq \eta'$; i.e., in period $t + 2$ the worker is in either job 1 with probability one or job 2 with probability one. We suspect that the corollary also holds for percentage wage changes when $\theta_L f(x + 2) < \eta' < \theta_H f(x+2)$ (i.e., there is a strictly positive probability that in period $t + 2$ the worker is in job 1 and a strictly positive probability the worker is in job 2). However, we have been unable to prove that the corollary holds for percentage wage changes in this case.

8. The reason that the possibility of demotion is important is as follows. In the symmetric-learning case when demotion is possible, the expected size of next period's wage increase is higher if the worker's current value for expected effective ability is close to either η' or η''. To see why, consider the case where the worker's value for expected effective ability exactly equals η'. In this case, increases in expected effective ability are valued at rate c_2 while decreases (which correspond to demotions) are valued at rate c_1. In effect, the lower slope c_1 offers partial insurance against wage decreases. This asymmetry, which is not a factor in the full-information model (because there workers never experience decreases in expected effective ability), means that the expected size of next period's wage increase is in fact larger than c_2 multiplied by the expected increase in expected effective ability. This feature of the symmetric-learning model is the reason that Corollary 1 does not necessarily hold when there is a positive probability of demotion in either period $t + 1$ or $t + 2$.

w_{it+1} is an increasing function of $w_{it+1} - w_{it}$. We also impose this no-demotion assumption in deriving another result below. Recall that the no-demotion case is empirically relevant: in the BGH data set only 0.3 percent of year-to-year job movements were demotions.

Now consider promotion rates. In the full-information model, if η' and $\eta'' - \eta'$ are both sufficiently large, then there is serial correlation in promotion rates. As opposed to the other results that hold in the full-information case, we have been unable to show that symmetric learning yields serial correlation in promotion rates. We can show a related but weaker result: given two workers earning the same wage just after promotion to job 2, the worker with less labor-market experience will on average be promoted to job 3 at a younger age. Similar to what is true under full information, we also know that the worker promoted earlier will have his or her expected effective ability grow more quickly, on average, than the expected effective ability for the worker promoted later, which suggests serial correlation in promotion rates. However, to prove serial correlation in promotion rates under symmetric learning requires a statement concerning how the distribution of expected effective ability evolves with experience, and this makes the problem intractable (at least for us).

Another prediction of the full-information model that matches the BGH evidence is the prediction that wage increases predict promotion. This prediction continues to hold in the symmetric-learning case in the following two ways. First, workers who receive higher wage increases after their first year of employment are more likely to be promoted at the first experience level at which promotion is possible (i.e., at experience level x'_H as defined in subsection III.B). Second, BGH find that workers who receive higher wage increases after their first year of employment are on average promoted to job 2 earlier. Our symmetric-learning model yields this result for parameterizations such that there are no demotions from job 2 to job 1.[9]

COROLLARY 2. Consider a worker in period t who has zero periods of labor-market experience and therefore is in job 1. Assume a

9. The corollary assumes a parameterization such that $x'_H < x'_L$. If $x'_H = x'_L$, then the model yields the degenerate result that all promotions occur at the same value for labor-market experience. Also, under symmetric learning it is possible that the worker will be promoted directly from job 1 to job 3. When this is the case, we define the value of the worker's labor-market experience upon promotion to job 2 to be the same as the value upon promotion to job 3.

parameterization such that $x'_H < x'_L$. (i) The probability of promotion at experience level x'_H is an increasing function of $w_{it+1} - w_{it}$. (ii) If demotion from job 2 to job 1 is impossible (i.e., for any x, $\theta_H f(x) > \eta'$ implies $\theta_L f(x+1) > \eta'$), then the expected value of the worker's labor-market experience when the worker is promoted to job 2 is a decreasing function of $w_{it+1} - w_{it}$.

Corollary 2 refers to promotion from job 1 to job 2. BGH also find that workers who receive higher wage increases after their first year in job 2 spend less time in job 2 before promotion to job 3. The symmetric-learning model also yields a result consistent with this finding for parameterizations such that there are no demotions from job 3 to job 2. Consider a worker in period t who has x periods of labor-market experience, has just been promoted to job 2, and is currently earning the wage w_{it}. If the worker is again in job 2 in period $t + 1$, then the expected value of the worker's labor-market experience when the worker is promoted to job 3 is a decreasing function of $w_{it+1} - w_{it}$.

The logic for why wage increases predict promotion is closely related to the above discussion concerning serial correlation in wage increases. As before, a large wage increase indicates that an increase in expected innate ability occurred, which means that on average the worker's expected effective ability will grow more quickly in the future. Since promotions occur when expected effective ability reaches certain absolute levels, on average a worker who experiences a large wage increase will need less time to reach the target level of expected effective ability needed for promotion.

We now consider predictions of the full-information case that are inconsistent with the BGH evidence. First, the full-information case predicts no demotions and no wage decreases, while the evidence indicates a positive frequency of both, although demotions are quite rare. In contrast to the full-information case, the symmetric-learning case yields predictions concerning demotions and wage decreases that are more consistent with the evidence.

COROLLARY 3. Let x^* be the smallest x such that $\theta_L f(x+1) < \theta_H f(x)$. Consider a worker in period t who has x periods of labor market experience. For $x < x^*$, there are no wage decreases and no demotions (i.e., $w_{it+1} - w_{it} \geq 0$ and the worker is not in a lower level job in $t + 1$ than in t). For $x \geq x^*$, there exists a positive frequency of wage decreases but fewer (possibly no) demotions. Any demotions that occur are associated with wage decreases.

In the full-information case, a worker's effective ability increases each period, so in equilibrium there are no wage decreases. With symmetric learning, wage decreases occur because a worker's expected effective ability does not necessarily increase every period: because of the learning process, a worker's expected innate ability can fall substantially from one period to the next; if this decrease is sufficiently large, it will dominate the increase in expected effective ability due to human-capital accumulation.[10]

Our argument explains not only a positive frequency of wage decreases, but also why wage decreases are a minority of the observations, although not rare, while demotions are very rare. Roughly half the workers will experience a decrease in expected innate ability, so if human-capital acquisition is positive but small, then many of these will also experience a decrease in expected effective ability. The result is that wage decreases are a minority of the observations, but may not be rare. To see why demotions might be very rare or nonexistent, consider all workers assigned to job 2 in period t, and suppose that $\eta'' - \eta'$ is large. All workers for whom expected effective ability falls will experience a wage decrease, but only those for whom the new value of expected effective ability is below η' will receive a demotion. If $\eta'' - \eta'$ is large, then few of the workers assigned to job 2 in period t will have values of η^e_{it} close to η', which implies that even fewer will have values of η^e_{it+1} below η'. Indeed, if $\theta_H f(x) > \eta'$ implies that $\theta_L f(x+1) > \eta'$, then there can be a positive frequency of wage decreases but no demotions.[11]

10. Two alternative reasons for real-wage decreases are aggregate shocks and person-specific productivity variations. The aggregate-shock explanation seems incomplete because the fraction of real-wage decreases in the BGH data set never fell below 5 percent and the median value of this fraction was 12 percent. Person-specific productivity variations, such as health shocks, would produce essentially the same qualitative predictions as our symmetric-learning model. Our instinct is that for prime-age managers continuously employed in a single firm, such productivity variations are unlikely to have sufficient size and frequency to match the evidence.

11. BGH find that on average demotions are associated with wage decreases, but that some demotions are in fact associated with wage increases. Our model does not explain why this would be the case. Two standard explanations for such a finding are coding error (this could be studied by examining job assignments in earlier and later periods) and compensating differentials (not modeled here). See

Another important aspect of the symmetric-learning equilibrium is the size of wage increases upon promotion. Recall that we found that in the full-information case, if human-capital acquisition is close to linear then the wage increase at promotion to job 2 is larger than the wage increase prior to promotion, consistent with the evidence. We also found, however, that under these conditions the wage increase the period *after* promotion to job 2 will be even larger than the wage increase at promotion, and this is not consistent with the evidence. The symmetric-learning model fixes this problem, as follows.

Consider workers promoted from job 1 to job 2 and assume again that human-capital acquisition is close to linear. As in the full-information case, the wage increase at promotion is larger than the wage increase prior to promotion, but there are now two reasons for this result. First, as before, increases in expected effective ability are valued at rate c_1 in job 1 and rate $c_2 > c_1$ in job 2. Second, in the symmetric-learning case there is a selection effect: among workers who are observationally equivalent at the beginning of a given period, those promoted at the end of that period had a larger increase in expected innate ability than did those not promoted. It is still the case that, after promotion to job 2, increases in expected effective ability are valued at rate c_2 rather than at a convex combination of c_1 and c_2, but the selection effect disappears after promotion to job 2. Hence, in the symmetric-learning case, the wage increase the period after a promotion can easily be smaller than the wage increase at promotion.

Although symmetric learning exhibits large wage increases upon promotion in a stronger form than does full information, under symmetric learning it is still the case that wage increases upon promotion explain only a fraction of the difference between average wages across levels of the job ladder. The logic is the same as under full information. In keeping with this finding, BGH also find that those promoted from one level of the job ladder to the next come disproportionately, but not exclusively, from the top of the lower job's wage distribution (and analogously arrive disproportionately, but not exclusively, at the bottom of the higher job's wage distribution). For example, averaging across all promotions from levels 1 through 4, 66 percent of workers promoted come

Bernhardt [1995] for another explanation in which wage increases are needed to stop demoted workers from being bid away by prospective employers.

from above the median of the lower job's wage distribution, and 72 percent arrive below the median of the higher job's wage distribution.[12] Corollary 4 proves a related result: the probability that a worker will be promoted is an increasing function of the worker's percentile in the lower job's wage distribution.

COROLLARY 4. Consider workers in period t who have x periods of labor-market experience and are in job j. Let $\pi(w)$ denote the probability that a worker in this group who is currently paid wage w is promoted for period $t + 1$. If $0 < \pi(w) < 1$, then $\pi(w)$ is increasing in w.

There are two factors that allow our model to capture the empirical findings concerning where workers are promoted from and where they are promoted to: heterogeneity in θ and learning about θ. The first of these factors exists even under full information, as follows. Consider all workers at level j. Since workers with low values for expected innate ability accumulate human capital more slowly than workers with high values for expected innate ability, the latter workers will have more average growth in expected effective ability. Hence, although most promoted workers will come from the top of the lower job's wage distribution, some may come from far below the top of the lower job's wage distribution. Analogously, most promoted workers will arrive near the bottom of the higher job's wage distribution, but some may arrive far above the bottom of the higher job's wage distribution.

Learning about θ produces a similar result. The idea is that, similar to what we argued above for human-capital accumulation, learning makes it possible for a worker's expected effective ability to increase substantially from one period to the next. Thus, there should be observations of workers being promoted from the low end of the lower job's wage distribution, but such observations should be rare. Again, an analogous argument holds for where workers arrive in the higher job's wage distribution.[13]

12. BGH also find a systematic relationship between job level and these percentages. That is, in most cases, the higher is the job from which promotion occurs, the higher is the percentage of promoted workers who come from above the median of the lower job's wage distribution, and the higher is the percentage who arrive below the median of the higher job's wage distribution. Specifically, for promotion from levels 1, 2, 3, and 4, the percentages of workers who come from above the median of the lower job's wage distribution are, respectively, 60, 67, 76, and 92, while the corresponding percentages for arriving below the median of the higher job's wage distribution are 65, 73, 88, and 81. As discussed in footnote 13, this pattern can also be explained by our model.

13. As discussed in footnote 12, BGH find that in most cases the higher is the job from which promotion occurs, the higher is the percentage of promoted workers

Overall, the symmetric-learning case does a better job than the full-information case of matching the BGH evidence: introducing symmetric learning leaves unchanged most of the predictions of the full-information case that match the evidence well, while improving several predictions that did not fit the evidence. In particular, under symmetric learning there can be a significant number of wage decreases but very few demotions. Also, symmetric learning does a better job matching the evidence concerning wage increases at promotions and the positions of promoted individuals in the wage distributions of the sending and receiving jobs.

V. Discussion

This section offers three discussions. First, we describe how our approach relates to evidence concerning performance evaluations. Second, we identify and discuss findings from the BGH study that are inconsistent with the predictions of our symmetric-learning model. Third, we consider how other models of wage dynamics inside firms relate to the evidence.

A. Performance Evaluations

In an influential pair of studies in the early 1980s, Medoff and Abraham [1980, 1981] (hereinafter MA) used the personnel records of three firms to explore the human-capital explanation

who come from above the median of the lower job's wage distribution, and the higher is the percentage who arrive below the median of the higher job's wage distribution. This pattern can be explained by the learning argument discussed above, because the incremental amount that is learned should fall as the worker ages. That is, because being higher on the job ladder will be positively correlated with age, our learning model predicts that both the percentage of workers promoted from the high end of the lower job's wage distribution and the percentage promoted to the low end of the higher job's wage distribution should rise as job level rises.

Related to BGH's results concerning where workers are promoted from and to in the job ladder's wage distributions, BGH also find that the wage distributions for adjacent job levels are overlapping. Although this finding conflicts with a literal interpretation of our model, a natural extension of our model captures this finding as follows. BGH's empirical analysis and our theoretical analysis treat the firm as if there were a single job ladder. But for most large firms there are multiple job ladders, such as those associated with different occupations. Given this, suppose we extended our model so that the firm has two job ladders but promotions occur only within ladders, not across. Suppose further that promotions occur only from the top of the wage distribution at one level of a ladder to the bottom of the wage distribution at the next level of that ladder. If the wage distributions for one ladder are shifted relative to those of another (say, because of compensating differentials or other occupational wage differences), then an analysis of the aggregate job ladder would yield overlapping wage distributions.

for why wages grow over a career. In their 1980 paper MA demonstrate that within-job wages are positively related to labor-market experience, but that performance evaluations are either unrelated or slightly negatively related to experience.[14] To test whether performance evaluations are a good measure of productivity, MA run two further regressions. First, they find that performance evaluations predict future promotions. Second, they (and also BGH) find that performance evaluations predict future wage increases.

In this subsection we build on a discussion in Harris and Holmstrom [1982] to offer an interpretation of performance evaluations that is consistent with both our theoretical model and MA's evidence, but that also preserves human-capital theory as the primary explanation for wage growth over a career. Harris and Holmstrom [p. 326] suggest that the MA findings might be explained if supervisors evaluate individuals relative to other individuals with the same labor-market experience. That is, in our terminology, performance evaluations measure expected innate ability rather than current expected productivity (i.e., expected effective ability).

Given this interpretation of performance evaluations, all the MA results described at the beginning of this subsection are easy to understand. In particular, within a job level, the average wage increases with experience, while the average performance evaluation falls with experience. In our model the wage is an increasing function of expected effective ability, which increases with experience because of human-capital accumulation. Further, performance evaluations within a job level fall with experience because workers with high expected innate ability are promoted into higher level jobs. Finally, performance evaluations predict both future promotions and future wage increases because both are positively related to expected innate ability.

Two other sources of evidence support our explanation of the MA findings. The first is a meta-analysis of 40 studies of age and job performance by Waldman and Avolio [1986]. Some of the studies contained direct measures of productivity, while others contained supervisory performance evaluations. Waldman and Avolio found that there was a significant positive relationship between direct measures of productivity and age, but a slight

14. MA also present results concerning firm seniority as well as labor-market experience. Because there is no turnover in equilibrium in our model, we do not address these MA findings.

negative relationship between supervisory performance evaluations and age. Thus, Waldman and Avolio provide clear evidence that performance evaluations are not unbiased measures of productivity. Furthermore, their results are consistent with MA's empirical findings, but do not imply that human-capital acquisition plays a minor role in wage growth. Rather, Waldman and Avolio's findings support our model's explanation for the MA evidence.

The second source of evidence supporting our explanation for the MA findings concerns a comparison between cross-sectional and longitudinal estimates from a single data set. Our model predicts that in a cross-sectional study within a given job level the selection effect concerning who earns promotions will cause the average wage to rise with experience but the average performance evaluation to fall. In contrast, consider a longitudinal study that follows a fixed set of workers who remain at the same job level over a specified period of time. Our model again predicts that the average wage will rise with experience, but now the selection effect will be smaller. That is, since the analysis follows the same workers over time, the only selection effect is that workers who remain at the same job level over a period of time will typically be those for whom learning led to more pessimistic beliefs concerning expected innate ability. Thus, in a longitudinal analysis within a job, we would expect a smaller negative relationship between the average performance evaluation and experience than in a cross-sectional analysis.

In their 1981 study Medoff and Abraham perform both cross-sectional and longitudinal regressions on the same data set and find results consistent with the above discussion (see also Gibbs [1995]). In the cross-sectional estimates they find that within a job level there is a significant positive relationship between the average wage and experience and a significant negative relationship between the average performance evaluation and experience. In the longitudinal estimates they find a significant positive relationship between the average wage and experience and a small negative relationship between the average performance evaluation and experience.

B. Empirical Findings We Do Not Capture

In Section IV and subsection V.A we focused on empirical findings in the BGH and MA studies that are consistent with our symmetric-learning approach. In this subsection we identify and discuss three findings from the BGH study that are not captured by our model. The first empirical finding inconsistent with our model involves nominal wage rigidity. Although BGH find a significant percentage of real wage decreases, nominal wage decreases are almost nonexistent in their sample—fewer than 200 observations in a sample of almost 70,000. (See Kahn [1997] and Card and Hyslop [1997] for other evidence that nominal wage decreases are rare.) Our model provides no rationale for why this should be the case, but an extension of our model does match some related evidence from BGH, as follows.

Suppose that one were to add to our model a constraint that firms could not offer nominal wage decreases and a hiring cost so that firms would not have an incentive to fire a worker whenever the worker's wage exceeded the worker's marginal product.[15] In a model with these two new features, if the information revealed about the worker in period t is very negative, then in period $t + 1$ the firm would be constrained to pay a wage higher than would be predicted by our original model. Further, since the actual wage in $t + 1$ is too high, the wage increase in $t + 2$ should be very low. BGH find evidence in this spirit: the probability of a zero nominal increase this year, given a zero nominal increase last year, is two to three times the unconditional probability.

The second finding inconsistent with our model involves cohort effects. BGH study cohorts who enter the firm's management at different dates and find that, even after controlling for (coarse) observable differences across cohorts, a cohort's average wage at entry is an important determinant of the cohort's average wage years after entry. Beaudry and DiNardo [1991] study cohort effects in panel data covering a large cross section of occupations and industries. Beaudry and DiNardo's first analysis is similar to BGH's: they find that the unemployment rate in a cohort's entry year 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 the cohort's entry year to affect current wages. They find that each of these three regressors is significant on its own but that the lowest

15. In our model of spot-market contracting, without a hiring cost, if firms cannot offer nominal wage decreases, they will fire a worker whenever the worker's wage exceeds the worker's marginal product. As discussed in footnote 2, a hiring cost means that a worker's employer earns a profit in the worker's last period of employment. As a result, the firm is willing to continue employing a worker even when the wage exceeds the worker's marginal product as long as it is by a sufficiently small amount.

unemployment rate since entry is the only significant regressor when all three are included.

It may be that the true structure of the BGH data is similar to Beaudry and DiNardo's finding. That is, the apparent influence of a cohort's entry wage on its current wage might be produced by the true influence of the highest entry wage paid since that cohort entered. Further, we conjecture that such a cohort effect would emerge from our model if we added nominal wage rigidity in the fashion discussed above. To keep things simple, suppose that there is no inflation, but add to our model aggregate shocks that move all workers' wages up or down. Nominal wage rigidity produces a ratchet effect: after a negative shock, wages will not fall; after a positive shock, competition will cause wages to rise. Hence, a cohort's current wage will depend on the strongest labor-market conditions since that cohort entered.

The last BGH finding that is inconsistent with the predictions of our model is the "green-card" effect. Consider the workers who are in job j at date t. The green-card effect refers to the idea that, for these workers, the expected wage increase, $E(w_{t+1}|w_t) - w_t$, is negatively related to the initial wage, w_t. (See Murphy [1991] for related evidence.) This result is not clearly inconsistent with our model. That is, because the human-capital acquisition function $f(x)$ is increasing and concave and workers higher in the wage distribution at a given job are likely to have more labor-market experience, our model predicts some negative correlation between $E(w_{t+1}|w_t) - w_t$ and w_t. Nevertheless, BGH find a strong green-card effect, and we feel it is unlikely that this result can be completely explained by the curvature of $f(x)$. The correct explanation may be that green-card effects are the result of administrative rules and procedures that move wages away from spot-market pay.

C. Other Theoretical Models

In the previous subsection we showed that our symmetric-learning model does not capture all the BGH findings. This prompts the question of whether another model might meet with more success. In this subsection we consider three alternative theories of wage dynamics inside firms: Bernhardt's [1995] model of asymmetric learning, Harris and Holmstrom's [1982] model of insurance, and Lazear and Rosen's [1981] tournament model. In each case, we assess the extent to which the alternative theory (or its natural extension) matches the BGH evidence.

Bernhardt's [1995] model is the closest of these alternatives to our model. In contrast to our symmetric-learning model, however, Bernhardt builds on Waldman's [1984b] model of asymmetric learning, in which only the current employer observes a worker's ability, but that firm's promotion decision signals information about the worker's ability to prospective employers. Bernhardt extends Waldman's model to more than two periods and allows workers to acquire general as well as specific human capital, although most of his analysis retains Waldman's assumption of only two job levels. In Bernhardt's model, promotions occur in bunches: first a group from the top of the ability distribution is promoted, and prospective employers draw a positive inference about those promoted and a negative inference about the remaining workers; then all workers acquire more specific capital, and a second group is promoted; and so on. From the perspective of matching the BGH evidence, the key feature of Bernhardt's model is that the only information prospective employers have about a worker is the worker's experience level and the date of the worker's promotion, if it has occurred.

Because of the signaling effect from promotions, Bernhardt's model matches the evidence that wage increases at promotions are large. And because of human-capital acquisition, wage increases at promotion may be small relative to the difference in average wages between levels of the job ladder. On the other hand, in Bernhardt's model a worker's wage depends solely on the worker's experience level and the age at which the worker was promoted (if that has occurred). But several BGH findings rely on heterogeneous wages being paid to workers who are identical in these two dimensions. For example, wage changes do not forecast promotions in Bernhardt's model because all workers who have the same experience level and who have not been promoted receive the same wage change. Similarly, negative real wage changes are possible, but occur in a fashion at odds with the data. Specifically, for a given experience level, all workers on the low-level job who are not promoted in a given period receive the same wage change, either positive or negative.

Harris and Holmstrom [1982] develop a model of learning and insurance, but do not incorporate job assignment or human-capital acquisition. As in our symmetric-learning model, firms learn about workers' abilities by observing noisy output signals. Unlike our model, workers are risk-averse and so are willing to pay for insurance. But a firm cannot offer full insurance, because a worker whose expected productivity rises above the wage guaranteed in a full-insurance contract can quit to accept a higher offer elsewhere. The optimal contract is thus downward rigid, guaranteeing a base wage below the worker's current expected productivity, but rising to meet the market if the worker's current or future outputs are sufficiently high.

Harris and Holmstrom describe how this model matches classic cross-sectional evidence on earnings profiles (e.g., Mincer [1974]). They also show that wages increase with experience even if productivity is constant, which provides a possible explanation for one of MA's major findings. But these downward-rigid contracts are inconsistent with the substantial frequency of negative real-wage changes. Furthermore, because there are no job assignments, the model does not address findings such as large wage increases at promotions or wage increases that forecast promotions. But Chiappori, Salanie, and Valentin [1999] offer an empirical implementation of the Harris-Holmstrom model in which wage increases occur only at promotions, so a worker is more likely to be promoted this period if the worker was more recently promoted to the current position. Chiappori, Salanie, and Valentin provide evidence consistent with this prediction.[16]

As a final alternative, we consider Lazear and Rosen's [1981] tournament model. This is a natural alternative, because it introduces incentives, which none of the foregoing models includes. We can interpret Lazear and Rosen's model in terms of two dates and two jobs. All workers are assigned to the first job and earn the same wage in the first period, but those promoted to the second job earn a higher wage in the second period. The larger the difference in the wages attached to the two jobs in the second period, the larger the incentive for all workers to work hard in the first period. This model was built to explain large wage increases upon promotion, but in its simplest form does not address other findings such as wage increases predicting promotion, a substantial frequency of real-wage decreases, and serial correlation in

16. A simple version of our symmetric-learning model also produces such a result for some parameterizations. Suppose that there are only three output realizations: high, intermediate, and low. Further, suppose that in the first period only workers who produce high output are promoted, while in the second period those who produce high output again are promoted a second time. Because those who produce high output twice are promoted twice, those who produce high output early but are promoted only once are a selected sample—the truly fast trackers are selected out, leaving only the lower tail of the distribution. For some parameterizations, this group can have poorer promotion prospects than those who are promoted for the first time in the second period.

wage increases. Meyer [1992] shows that a repeated-tournament model can capture serial correlation in promotions (and hence possibly in wage increases). It remains an open question whether further extensions of the tournament model could capture other BGH findings.

In summary, we believe our symmetric-learning model does a better job than the existing alternatives of capturing the BGH findings. This is not to say that there is not an extension of one of these alternatives or a completely new model that would do even better than the framework we have developed. Our view is simply that this framework provides the best current explanation for the broad pattern of findings in the BGH and MA studies.[17]

VI. Conclusion

Recently, several empirical studies of careers in organizations have found results that are inconsistent with theoretical models based on any single factor such as human-capital acquisition or learning. In this paper we developed a model that integrates job assignment, on-the-job human-capital acquisition, and learning. We demonstrated that a model that combines these elements captures many of the findings in this recent literature. Our conclusion is not that our symmetric-learning model provides a perfectly accurate representation of careers in organizations. Rather, we find the results encouraging, and feel that our model offers a basic framework upon which to build more accurate models.

We have interpreted our model as applying to wage and promotion dynamics within firms. But our model assumes that wages are determined by spot-market competition and that an infinitesimal moving cost eliminates mobility between firms (although see footnote 2). Thus, if we added exogenous reasons for

17. Kwon's [1998] promising paper begins to capture the BGH findings using an approach different from ours. In Kwon's model, a worker must decide each period whether to invest in human capital. The optimal contract has several features that match the BGH evidence. For example, there are large wage increases upon promotion, and the wage distributions for adjacent job levels are overlapping (see footnote 13). One problem with Kwon's approach, however, is that a worker's wage does not change from period to period unless a promotion occurs. Thus, the model does not capture some of the prominent BGH findings such as that wage changes are serially correlated or that workers who receive larger wage increases early in their stay at one level of the job ladder are promoted more quickly to the next level.

mobility to the model, all our results would continue to hold but the model would apply to wage and promotion dynamics in the labor market as a whole rather than only inside firms. The recent evidence developed by Baker [1997] is consistent with this idea: as discussed in Section II, Baker finds serial correlation in wage changes in a twenty-year panel of adult U. S. males, as would be predicted by this variant of our model. Unfortunately, many of the other predictions of this variant of our model involve promotions and so would be difficult to test using currently available data.

One direction in which the analysis might fruitfully be extended would be to incorporate an element of asymmetric learning. In this paper we assumed that all information concerning a worker's ability is public knowledge, so firms are always equally informed about a worker's ability. A number of authors have argued that this assumption is unrealistic: most employment relationships are characterized by an element of asymmetric learning, so that a worker's current employer has better information concerning the worker's ability than do prospective employers. As discussed above, Bernhardt [1995] develops a model similar to ours but with asymmetric learning rather than our symmetric learning. The extension we envision would blend the two, producing several additional results such as the following. First, as discussed above, asymmetric learning introduces an additional reason why promotions are associated with large wage increases. In addition, asymmetric learning means that a firm should be more certain concerning the ability levels of workers with high seniority, and this should translate into predictable differences in the future outcomes of newly hired workers versus those with long-term attachments. In fact, BGH find such a result: for workers entering level 3 between 1970 and 1979, workers who have been at the firm longer have a lower variance concerning the highest level of the firm attained through subsequent promotions.

In summary, we feel the analysis in this paper demonstrates that job assignment, on-the-job human-capital acquisition, and symmetric learning are important building blocks for constructing models of careers in organizations. However, incorporating asymmetric learning and other elements such as incentives will be important for extending our framework so that it more accurately matches the wage and promotion dynamics inside firms.

APPENDIX

Proof of Proposition 1

A worker with effective ability η has expected output $Ey_j = d_j + c_j\eta$ in job j. The efficient task assignment is therefore job 1 if $\eta \leq \eta'$, job 2 if $\eta' < \eta \leq \eta''$, and job 3 if $\eta'' < \eta$. Competition among firms yields both efficient task assignment and wages equal to expected output: $w = d_j + c_j\eta$ for the efficient job j.

QED

Proof of Proposition 2

The model has been constructed so that the argument in the symmetric-learning case can parallel the argument in the full-information case: wages again equal expected output each period, now given the observed history from prior periods. Downward-rigid contracts such as those in Harris and Holmstrom [1982] are feasible here but have no benefits because workers are risk-neutral. We compute a worker's expected effective ability η^e in (3) and then the worker's expected output in job j as $Ey_j = d_j + c_j\eta^e$. The linearity of the production function (2) is key here: without linearity, expected output would not equal the output of a worker known to have ability η^e. Finally, the fact that (2) reads $d_j + c_j(\eta_{it} + \epsilon_{ijt})$ rather than $d_j + c_j\eta_{it} + \epsilon_{ijt}$ means that the signal about ability that can be extracted from output—namely, $z_{it} = (y_{ijt} - d_j)/c_j = \eta_{it} + \epsilon_{ijt}$—does not vary in its signal-to-noise ratio as a function of j. Thus, there is no way to use task assignment to change the speed of learning about ability, so task assignment is determined by current productive efficiency (i.e., maximizing expected output this period), which in turn is solely a function of the worker's current expected effective ability.

QED

For expositional simplicity, the proofs of Corollaries 1 through 4 are presented in the following order: 3, 4, 1, 2.

Proof of Corollary 3

We first show that if $\theta_L f(x+1) < \theta_H f(x)$ then $\theta_L f(x+2) < \theta_H f(x+1)$. We do this by showing that $f(x)/f(x+1)$ increases in x, so that $f(x)/f(x+1) < f(x+1)/f(x+2)$, so that $\theta_L/\theta_H < f(x)/f(x+1)$ implies that $\theta_L/\theta_H < f(x+1)/f(x+2)$. Taking the derivative of $f(x)/f(x+1)$ shows that this fraction is increasing in x if $f'(x)f(x+1) > f(x)f'(x+1)$, which holds because $f(x)$ is increasing and concave. Therefore, $\theta_L f(x+1) < \theta_H f(x)$ for any $x > x^*$.

For $x < x^*$, we know that $\theta_L f(x+1) \geq \theta_H f(x)$, so any worker's expected effective ability must be higher when that worker has $x+1$ periods of prior experience than when he or she has x. Consequently, there can be no wage decreases or demotions.

For $x \geq x^*$, we show that there is a positive frequency of wage decreases by showing that beliefs about innate ability can move from arbitrarily close to θ_H to arbitrarily close to θ_L in one period. Let z^x denote the normalized output history $(z_{it-x}, \ldots, z_{it-1})$, where we drop the subscript i on z^x for simplicity. Let $p =$ prob $(\theta = \theta_H | z^x)$ be the probability that the worker has high innate ability given z^x. By Bayes' rule,

(A1)

$$\text{prob}\,(\theta = \theta_H | z^x, z_{it}) = \frac{ph[z_{it} - \theta_H f(x)]}{ph[z_{it} - \theta_H f(x)] + (1-p)h[z_{it} - \theta_L\ f(x)]},$$

where $h[\cdot]$ is the density of ϵ_{ijt} in (2), normal with mean 0 and variance σ^2. But

$$\text{(A2)} \quad \frac{h[z_{it} - \theta_L\ f(x)]}{h[z_{it} - \theta_H f(x)]} = \exp\left[-\frac{1}{2\sigma^2}\{[z_{it} - \theta_L f(x)]^2 - [z_{it} - \theta_H f(x)]^2\}\right]$$

$$= \exp\left[-\frac{1}{2\sigma^2}\{-2z_{it} f(x)(\theta_L - \theta_H) + f(x)^2(\theta_L^2 - \theta_H^2)\}\right],$$

which is monotonically decreasing in z_{it}, approaching 0 as $z_{it} \to \infty$ and approaching infinity as $z_{it} \to -\infty$. Therefore, prob $(\theta = \theta_H | z^x, z_{it})$ approaches one as $z_{it} \to \infty$ and zero as $z_{it} \to -\infty$. That is, a sufficiently strong signal can move the market's belief arbitrarily far.

Because a sufficiently strong signal can move the market's belief arbitrarily far and $\theta_L f(x+1) < \theta_H f(x)$, wage decreases occur with positive probability. If x is such that $\theta_L f(x+1) < \eta' < \theta_H f(x)$ or $\theta_L f(x+1) < \eta'' < \theta_H f(x)$, then demotions also occur with positive probability. Such demotions entail wage decreases: because $\eta > \eta'$ implies that $d_2 + c_2\eta > d_1 + c_1\eta$, we have (with some abuse of notation) $w_{it} = d_2 + c_2\eta_x^e > d_1 + c_1\eta_x^e > d_1 + c_1\eta_{x+1}^e = w_{x+1}$, and similarly for $\eta > \eta''$. Of course, the reverse is not true: because there is learning, there may often be wage decreases that occur without demotions.

QED

Proof of Corollary 4

For simplicity, we consider workers in job 1; analogous arguments apply to job 2. For workers with experience x, the assumption in the corollary that $0 < \pi(w) < 1$ can be restated as $\theta_L f(x+1) < \eta' < \theta_H f(x+1)$. That is, a worker with $x + 1$ periods of experience remains in job 1 if the belief about her ability is sufficiently pessimistic, but is promoted to job 2 (or 3) if this belief if sufficiently optimistic.

Recall that the prior belief prob $(\theta = \theta_H | z^x)$ at the beginning of period t is denoted by p and that the updated belief prob $(\theta = \theta_H | z^x, z_{it})$ at the beginning of period $t + 1$ is given in (A1). For simplicity, we will sometimes write the updated belief as q. Because the worker is assigned to job 1 in period t after x periods of experience, we know that $[p\theta_H + (1-p)\theta_L] f(x) < \eta'$. To be promoted to job 2 (or 3) for period $t + 1$, the worker's performance in period t, z_{it}, must be sufficiently high that $q \geq p^*_{x+1}$, where $[p^*_{x+1}\theta_H + (1 - p^*_{x+1})\theta_L] f(x+1) = \eta'$. That is, given p, z_{it} must satisfy

$$\text{(A3)} \qquad \frac{p}{p + (1-p)h[z_{it} - \theta_L f(x)]/h[z_{it} - \theta_H f(x)]} \geq p^*_{x+1}.$$

From (A2), $h[z_{it} - \theta_L f(x)]/h[z_{it} - \theta_H f(x)]$ is monotonically decreasing in z_{it}. Hence, given p, there exists a critical value $z^*_{x+1}(p)$ such that (A3) holds if and only if $z_{it} \geq z^*_{x+1}(p)$. That is, given past performance, there exists a critical value of current performance above which promotion occurs. From (A3), $z^*_{x+1}(p)$ decreases with p. That is, the critical value of current performance above which promotion occurs is lower if past performance has produced a more optimistic belief about innate ability. Furthermore, the conditional probability that z_{it} exceeds an arbitrary cutoff z^* increases in p, because

$$\begin{aligned}
\text{(A4)} \quad & \text{prob}(z_{it} \geq z^* | z^x) \\
&= \text{prob}(z_{it} \geq z^* | z^x, \theta_H)\, p + \text{prob}(z_{it} \geq z^* | z^x, \theta_L)(1-p) \\
&= \text{prob}[\epsilon_{ijt} \geq z^* - \theta_H f(x)]\, p + \text{prob}[\epsilon_{ijt} \geq z^* - \theta_L f(x)](1-p) \\
&= \text{prob}[\epsilon_{ijt} \geq z^* - \theta_L f(x)] \\
&\quad + p\{\text{prob}[\epsilon_{ijt} \geq z^* - \theta_H f(x)] - \text{prob}[\epsilon_{ijt} \geq z^* - \theta_L f(x)]\},
\end{aligned}$$

where $\text{prob}[\epsilon_{ijt} \geq z^* - \theta_H f(x)] - \text{prob}[\epsilon_{ijt} \geq z^* - \theta_L f(x)] > 0$. Because the critical value decreases in p and the probability that

z_{it} exceeds an arbitrary value increases in p, the probability of promotion increases in p.

The prior belief p is monotonically related to the current wage: $w = d_1 + c_1[p\theta_H + (1 - p)\theta_L]f(x)$. Thus, the critical value of z_{it} such that promotion occurs decreases in w. Furthermore, the probability that z_{it} exceeds an arbitrary cutoff increases in w. For both of these reasons, the probability of promotion increases in w.

QED

Proof of Corollary 1

The corollary assumes that the worker is in job 1 for periods t and $t + 1$, but is silent about period $t + 2$. We break the proof into three cases: (1) $\theta_H f(x + 2) \leq \eta'$, so that promotion is impossible after period $t + 1$; (2) $\theta_L f(x + 2) \geq \eta'$, so that promotion is guaranteed after period $t + 1$; and (3) $\theta_L f(x + 2) < \eta' < \theta_H f(x + 2)$, so that promotion is possible but not guaranteed after period $t + 1$.

In all three cases, we continue to use the notation above: $p = \text{prob}(\theta = \theta_H | z^x)$, so that $w_{it} = d_1 + c_1[p\theta_H + (1 - p)\theta_L] f(x)$, and $q = \text{prob}(\theta = \theta_H | z^x, z_{it})$, so that $w_{i,t+1} = d_1 + c_1[q\theta_H + (1 - q)\theta_L] f(x + 1)$. Thus,

$$\text{(A5)} \quad w_{i,t+1} - w_{it} = c_1[q\theta_H + (1 - q)\theta_L] f(x + 1) - c_1[p\theta_H + (1 - p)\theta_L] f(x)$$

increases in q given p. In each case, therefore, we show that $E(w_{i,t+2} - w_{i,t+1} | p,q)$ increases in q given p, and hence that $E(w_{i,t+2} - w_{i,t+1} | w_{it}, w_{i,t+1} - w_{it})$ increases in $w_{i,t+1} - w_{it}$ given w_{it}. To express the wage $w_{i,t+2}$, we introduce one last piece of notation: $r = \text{prob}(\theta = \theta_H | z^x, z_{it}, z_{i,t+1})$.

Case (1) is simple, because the worker is certain to be in job 1 for period $t + 2$, so $w_{i,t+2} = d_1 + c_1[r\theta_H + (1 - r)\theta_L] f(x + 2)$. But beliefs are a martingale, so $E(r|q) = q$. Thus, as of the beginning of period $t + 1$, the worker's expected wage increase from $t + 1$ to $t + 2$ is simply the slope c_1 times the expected increase in the worker's effective ability due to human-capital acquisition:

$$\text{(A6)} \quad E(w_{i,t+2} - w_{i,t+1} | w_{it}, w_{i,t+1} - w_{it}) = c_1[q\theta_H + (1 - q)\theta_L][f(x + 2) - f(x + 1)],$$

which increases in q.

Case (2) involves one more step than case (1): the worker's expected increase in effective ability from $t + 1$ to $t + 2$ is again $[q\theta_H + (1 - q)\theta_L][f(x + 2) - f(x + 1)]$, but now part of this increase is rewarded at c_1 and the rest at the higher slope c_2. But both of these effects favor workers with larger wage increases: the expected increase in effective ability is larger, and more of this increase is valued at the higher slope c_2. To see this, let $\eta^e_{t+1} = [q\theta_H + (1 - q)\theta_L]f(x + 1)$ denote the worker's expected effective ability at the beginning of period $t + 1$. We know that $\eta^e_{t+1} < \eta'$ because the worker remains in job 1 for period $t + 1$. Thus, the expected increase in effective ability from $t + 1$ to $t + 2$ can be written as all being rewarded at slope c_2 except for the part from η^e_{t+1} up to η', which is rewarded at c_1. That is,

$$\text{(A7)} \quad E(w_{i,t+2} - w_{i,t+1} | w_{it}, w_{i,t+1} - w_{it})$$
$$= c_2[q\theta_H + (1 - q)\theta_L][f(x + 2) - f(x + 1)] - (c_2 - c_1)(\eta' - \eta^e_{t+1}).$$

The first term increases in q, as in case (1), and $\eta' - \eta^e_{t+1}$ decreases in q.

Case (3) is another step more complicated than case (2), because now it is not certain that the worker will be promoted for period $t + 2$. In this case, it is convenient to rearrange the terms compared with (A7): the expected wage increase can be written as the entire expected increase in effective ability rewarded at slope c_1, plus any increase above η' rewarded at the incremental slope $(c_2 - c_1)$. That is,

$$\text{(A8)} \quad E(w_{i,t+2} - w_{i,t+1} | w_{it}, w_{i,t+1} - w_{it})$$
$$= c_1[q\theta_H + (1 - q)\theta_L][f(x + 2) - f(x + 1)]$$
$$+ (c_2 - c_1)E[g(\eta_{t+2}) | w_{it}, w_{i,t+1} - w_{it}),$$

where $\eta_{t+2} = [r\theta_H + (1 - r)\theta_L]f(x + 2)$ is the worker's expected effective ability at the beginning of period $t + 2$ and $g(\eta_{t+2}) = \max\{0, \eta_{t+2} - \eta'\}$. As in cases (1) and (2), the first term increases in q. It therefore remains to show that $E(g(\eta_{t+2}) | w_{it}, w_{i,t+1} - w_{it}) = E(g(\eta_{t+2}) | p, q)$ increases in q given p.

Because $g(\eta_{t+2})$ is an increasing function, it suffices to show first-order stochastic dominance: for any η^* between $\theta_L f(x + 2)$ and $\theta_H f(x + 2)$, $\text{prob}(\eta_{t+2} \geq \eta^* | p,q)$ increases in q. Because $\eta_{t+2} = [r\theta_H + (1 - r)\theta_L]f(x + 2)$, it suffices to show that for any r^* in $[0,1]$, $\text{prob}(r \geq r^* | p,q)$ increases in q. The argument is analogous to Corollary 4 but advanced one period, as follows.

Given q, there exists $z^*_{x+2}(q)$ such that $r = \text{prob}(\theta = \theta_H | z^x, z_{it}, z_{i,t+1}) \geq r^*$ if and only if $z_{i,t+1} \geq z^*_{x+2}(q)$. The critical value $z^*_{x+2}(q)$ decreases in q. Furthermore, the conditional probability that $z_{i,t+1}$ exceeds an arbitrary cutoff z^* increases in q. For these two reasons, the probability that $z_{i,t+1}$ exceeds the critical value increases in q.

QED

Proof of Corollary 2

The worker has zero periods of experience, so $w_{it} = d_1 + c_1[p_0\theta_H + (1 - p_0)\theta_L] f(0)$. We assumed in subsection III.A that $\theta_H f(1) < \eta'$, so every worker spends at least two periods in job 1. Thus, the next wage will be $w_{i,t+1} = d_1 + c_1[q_0\theta_H + (1 - q_0)\theta_L] f(1)$, where q_0 denotes $\text{prob}(\theta = \theta_H | z_{it})$. The wage change $w_{i,t+1} - w_{it}$ therefore increases in q_0 given p_0.

Because of human-capital acquisition, the worker will eventually be promoted to job 2. The lowest experience at which promotion can occur is x'_H (defined in subsection III.B as the lowest value of x such that $\theta_H f(x) > \eta'$). Recall from Corollary 4 that a worker with experience x'_H will not be in job 1 if the belief $p = \text{prob}(\theta = \theta_H | z^x)$ at $x = x'_H$ exceeds the critical value p^* that solves $[p^*\theta_H + (1 - p^*)\theta_L] f(x'_H) = \eta'$. Part (i) of the corollary claims that the probability that a worker with experience x'_H will not be in job 1 is an increasing function of the first wage change, $w_{i,t+1} - w_{it}$. So it suffices to show that the probability that p exceeds p^* is increasing in q_0. But this is a special case of Corollary 4, extended to allow an arbitrary number of periods between the realization of q_0 after one period of experience and the assessment of whether $p \geq p^*$ after x'_H periods of experience.

Part (ii) invokes a new assumption: that demotions are impossible. Under this assumption, the workers who are not in job 1 at any given date are precisely those who have been promoted by that date. That is, no workers are promoted earlier but then reassigned to job 1 on the date in question. Recall that a worker with experience x is not in job 1 if $p = \text{prob}(\theta = \theta_H | z^x) \geq p^*_x$, where $[p^*_x\theta_H + (1 - p^*_x)\theta_L] f(x) = \eta'$. Thus, the probability that a worker with experience x and first wage change $w_{i,t+1} - w_{it}$ is not in job 1 is $\text{prob}(p \geq p^*_x | q_0)$. This probability increases in q_0, for the same reasons as given in part (i). Thus, considering all the experience levels x from x'_H to x'_L, the cumulative probability of promotion $\text{prob}(p \geq p^*_x | q_0)$ is higher at every x for a higher value of q_0, in the sense of first-order stochastic dominance. Hence, the expected date of promotion is lower for a higher value of q_0.

QED

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CORNELL UNIVERSITY

REFERENCES

Abowd, J., and D. Card, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica,* LVII (1989), 411–445.
Baker, G., M. Gibbs, and B. Holmstrom, "The Wage Policy of a Firm," *Quarterly Journal of Economics,* CIX (1994a), 881–919.
Baker, G., M. Gibbs, and B. Holmstrom, "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics,* CIX (1994b), 921–955.
Baker, G., M. Jensen, and K. J. Murphy, "Compensation and Incentives: Practice vs. Theory," *Journal of Finance,* XLIII (1988), 593–616.
Baker, M., "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics,* XV (1997), 338–375.
Beaudry, P., and J. DiNardo, "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Microdata," *Journal of Political Economy,* XCIX (1991), 665–688.
Becker, G., *Human Capital* (New York, NY: NBER, 1964).
Ben-Porath, Y., "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy,* LXXV (1967), 352–365.
Bernhardt, D., "Strategic Promotion and Compensation," *Review of Economic Studies,* LXII (1995), 315–339.
Bernhardt, D., and D. Scoones, "Promotion, Turnover, and Preemptive Wage Offers," *American Economic Review,* LXXXIII (1993), 771–791.
Bruderl, J., A. Diekmann, and P. Preisendorfer, "Patterns of Intraorganizational Mobility: Tournament Models, Path Dependency, and Early Promotion Effects," *Social Science Research,* XX (1991), 197–216.
Card, D., and D. Hyslop, "Does Inflation 'Grease the Wheels of the Labor Market'?" in *Reducing Inflation: Motivation and Strategy,* C. Romer and D. Romer, eds. (Chicago, IL: University of Chicago Press, 1997), pp. 71–114.
Carmichael, L., "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics,* XIV (1983), 251–258.
Chiappori, P.-A., B. Salanie, and J. Valentin, "Insurance, Learning, and Career Profiles: An Empirical Test," *Journal of Political Economy,* CVII (1999), 731–760.
Demougin, D., and A. Siow, "Careers in Ongoing Hierarchies," *American Economic Review,* LXXXIV (1994), 1261–1277.
Doeringer, P., and M. Piore, *Internal Labor Markets and Manpower Analysis* (Lexington, MA: Heath Lexington Books, 1971).
Farber, H., and R. Gibbons, "Learning and Wage Dynamics," *Quarterly Journal of Economics,* CXI (1996), 1007–1047.
Gerhart, B., and G. Milkovich, "Salaries, Salary Growth, and Promotions of Men and Women in a Large, Private Firm," in *Pay Equity: Empirical Inquiries,* R. Michael, H. Hartmann, and B. O'Farrell, eds. (Washington, DC: National Academy Press, 1989), pp. 23–43.
Gibbons, R., "Incentives and Careers in Organizations," in *Advances in Economic Theory and Econometrics,* Vol. 2, D. Kreps and K. Wallis, eds. (New York, NY: Cambridge University Press, 1997), pp. 1–37.
Gibbons, R., and L. Katz, "Layoffs and Lemons," *Journal of Labor Economics,* IX (1991), 351–380.
Gibbons, R., and L. Katz, "Does Unmeasured Ability Explain Inter-Industry Wage Differences?" *Review of Economic Studies,* LIX (1992), 515–535.

Gibbons, R., and M. Waldman, “Careers in Organizations: Theory and Evidence,” in *Handbook of Labor Economics,* Vol. 3, O. Ashenfelter and D. Card, eds. (Amsterdam, The Netherlands: North-Holland, 1999), pp. 2373–2437.
Gibbs, M., “Incentive Compensation in a Corporate Hierarchy,” *Journal of Accounting and Economics,* XIX (1995), 247–277.
Greenwald, B., “Adverse Selection in the Labor Market,” *Review of Economic Studies,* LIII (1986), 325–347.
Harris, M., and B. Holmstrom, “A Theory of Wage Dynamics,” *Review of Economic Studies,* XLIX (1982), 315–333.
Hashimoto, M., “Firm-Specific Human Capital as a Shared Investment,” *American Economic Review,* LXXI (1981), 475–481.
Hause, J., “The Fine Structure of Earnings and the on-the-Job Training Hypothesis,” *Econometrica,* XLVIII (1980), 1013–1030.
Holmstrom, B., “Managerial Incentive Problems—A Dynamic Perspective,” in *Essays in Economics and Management in Honour of Lars Wahlbeck* (Helsinki, Finland: Swenska Handelshogkolan, 1982), pp. 209–230.
Jovanovic, B., and Y. Nyarko, “Stepping Stone Mobility,” *Carnegie-Rochester Series for Public Policy,* XLVI (1997), 289–337.
Kahn, C., and G. Huberman, “Two-Sided Uncertainty and Up-or-Out Contracts,” *Journal of Labor Economics,* VI (1988), 423–444.
Kahn, S., “Evidence of Nominal Wage Stickiness from Microdata,” *American Economic Review*, LXXXVII (1997), 993–1008.
Kwon, I., “Incentives, Wages, and Promotions: Theory and Evidence,” mimeo, Harvard University, 1998.
Lazear, E., “The Job as a Concept,” in *Performance Measurement, Evaluations, and Incentives,* W. Bruns, ed. (Boston, MA: Harvard University Press, 1992), pp. 183–215.
Lazear, E., and S. Rosen, “Rank-Order Tournaments as Optimum Labor Contracts,” *Journal of Political Economy,* LXXXIX (1981), 841–864.
Lillard, L., and Y. Weiss, “Components of Variation in Panel Data: American Scientists 1960–1970,” *Econometrica,* XLVII (1979), 437–454.
MacDonald, G. M., “A Market Equilibrium Theory of Job Assignment and Sequential Accumulation of Information,” *American Economic Review,* LXXII (1982), 1038–1055.
MacDonald, G. M., and J. R. Markusen, “A Rehabilitation of Absolute Advantage,” *Journal of Political Economy,* XCIII (1985), 277–297.
MacLeod, B., and J. Malcomson, “Reputation and Hierarchy in Dynamic Models of Employment,” *Journal of Political Economy,* XCVI (1988), 832–854.
Main, B., C. O’Reilly, and J. Wade, “Top Executive Pay: Tournaments or Teamwork?” *Journal of Labor Economics*, XI (1993), 606–628.
McCue, K., “Promotions and Wage Growth,” *Journal of Labor Economics,* XIV (1996), 175–209.
McLaughlin, K., “Rigid Wages,” *Journal of Monetary Economics,* XXXIV (1994), 383–414.
Medoff, J., and K. Abraham, “Experience, Performance, and Earnings,” *Quarterly Journal of Economics,* XCV (1980), 703–736.
Medoff, J., and K. Abraham, “Are Those Paid More Really More Productive?” *Journal of Human Resources,* XVI (1981), 186–216.
Meyer, M. A., “Biased Contests and Moral Hazard: Implications for Career Profiles,” *Annales d’Economie et de Statistique,* XXV (1992), 165–187.
Mincer, J., *Schooling, Experience, and Earnings* (New York, NY: NBER, 1974).
Murphy, K. J., “Corporate Performance and Managerial Remuneration: An Empirical Analysis,” *Journal of Accounting and Economics,* VII (1985), 11–42.
——, “Incentives, Learning, and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts,” *Rand Journal of Economics,* XVII (1986), 59–76.
——, “Merck & Co., Inc. (A),” Harvard Business School Case #9-491-005, 1991.
Podolny, J., and J. Baron, “Resources and Relationships: Social Networks and Mobility in the Workplace,” *American Sociological Review,* LXII (1997), 673–693.
Prendergast, C., “The Role of Promotion in Inducing Specific Human Capital Acquisition,” *Quarterly Journal of Economics,* CVIII (1993), 523–534.

Ricart i Costa, J., "Managerial Task Assignment and Promotions," *Econometrica,* LVI (1988), 449–466.
Rosen, S., "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics,* XIII (1982), 311–323.
Rosenbaum, J., *Career Mobility in a Corporate Hierarchy* (New York, NY: Academic Press, 1984).
Ross, S., P. Taubman, and M. Wachter, "Learning by Observing and the Distribution of Wages," in *Studies in Labor Markets,* S. Rosen, ed. (Chicago, IL: University of Chicago Press, 1981).
Sattinger, M., "Comparative Advantage and the Distribution of Earnings and Abilities," *Econometrica,* XLIII (1975), 455–468.
____, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature,* XXXI (1993), 831–880.
Topel, R., "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy,* XCIX (1991), 145–176.
Topel, R., and R. Ward, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics,* CVII (1992), 439–480.
Waldman, D., and B. Avolio, "A Meta-Analysis of Age Differences in Job Performance," *Journal of Applied Psychology,* LXXI (1986), 33–38.
Waldman, M., "Worker Allocation, Hierarchies, and the Wage Distribution," *Review of Economic Studies,* LI (1984a), 95–109.
____, "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics,* XV (1984b), 255–267.
____, "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics,* VIII (1990), 230–250.