Internal Labor Markets in Equilibrium[*]

Timothy N. Bond[†]
Purdue University

November 27, 2012

Abstract

Traditional models of promotion have difficulty explaining why many firms do not favor their own workers for advancement. I develop a new model which seeks to explain this phenomenon. My model generates an equilibrium in which some firms, but not all, commit to promote ex-ante. These "promotion" firms are able to attract higher quality employees to the lower ranks of the firm, which in turn makes them more likely to get skilled workers in upper level jobs. Non-promotion firms benefit from paying lower wages. This divergence in strategies is due to the scarcity of high quality workers. My model generates several testable predictions: first, workers at entry-level jobs will be paid higher wages at promotion firms, both because they are better workers and because they capture some of the quasi-rents that promotion firms earn after workers are promoted. Second, the observed return to tenure in a cross-section of entry-level job workers will be higher at non-promotion firms because the best workers in all firms will advance over time to higher jobs, either inside or outside the firm. Third, there will be no wage differences between workers at promotion and non-promotion firms among the most senior employees at entry-level jobs, since only low ability workers will remain at entry-level jobs at any firm. Finally, workers at supervisorial jobs will earn higher wages at promotion than non-promotion firms, because committing to promote yields better candidates for these jobs in equilibrium. I confirm these predictions empirically using a matched employer-employee data set from the UK.

---

[*] I am extremely grateful to Kevin Lang, Michael Manove, and Claudia Olivetti for their continuing advice and support. I would also like to thank Marric Buessing, Andrew Ellis, Daniele Paserman, Dana Rotz, Laura Salisbury, Johannes Schmieder, and seminar participants at Boston College, Boston University, Carleton University, Miami University, Purdue University, the 2012 Southern Economic Association Annual Meetings, the University of Adelaide, and Università Bocconi for their helpful comments. I acknowledge the Department of Trade and Industry, the National Institute of Economic and Social Research, the Advisory, Conciliation, and Arbitrary Service, and the Policy Studies Institute as the originators of the 2004 Worker Employee Relations Survey data and the UK Data Archive as the distributor of the data. These organizations bear no responsibility for the analysis of this data herein. All errors are my own.

[†] Department of Economics, Krannert School of Management, Purdue University, 403 W. State Street, West Lafayette, IN, 47907; email: tnbond@purdue.edu

1 Introduction

Typical models of promotion focus on the provision of incentives within the firm. In the classic tournament setting of Lazear and Rosen (1981), workers are motivated to exert the optimal effort by a prize that goes to the worker who produces the highest output. However, these models have difficulty explaining the empirical finding that many firms do not favor their own workers for advancement. In my data set, 71% of workers are employed by a firm that gives no preference to internal candidates, in spite of the fact that such favoritism is necessary in order to effectively incentivize employees in a tournament setting (Chan, 1996).

I develop a model in which some firms choose to promote internally because this attracts better quality candidates into entry-level jobs. Promotion also allows these firms to avoid hiring higher-level workers from a frictional market. This synergy drives the formation of internal labor markets given available talent. The fact that talent is scarce limits the viability of this strategy. In equilibrium the number of firms that make this commitment is determined by the number of workers in the labor pool who appear promising *ex ante*. This model produces several predictions about how the labor force and wages of promotion and non-promotion firms differ, which I test and confirm in the empirical section.

The intuition behind the model is quite simple. Imagine a world where talent is both scarce and valuable. Because firms can extract rents from workers, they would like to hire the most talented employees and place them in their most productive positions. However, due to the scarcity of talent, some firms will be forced to fill their best positions with lower quality workers. By "locking in" a promising young worker and guaranteeing promotion regardless of whether he actually develops into a highly skilled worker, the firm can increase the probability of eventually having a high-quality worker at an upper-level position.[1] The worker benefits by getting a chance to work at a higher wage position regardless of whether he develops into a premium talent. Of course, many firms would like to adopt this strategy and competition will bid up the wages for promising young workers. Given this, some firms choose to wait and attempt to hire workers who, although initially less promising, turned out to be high quality.

---

[1] This intuition is similar to that of Li and Rosen (1998), though their model deals with risk aversion in incomplete markets.

I model an overlapping generations economy in which there are two types of workers who differ in their productivity across all jobs. Workers' true types are unknown, but a fraction of these workers give off a publicly observable signal that makes them more likely to be the good (i.e. more productive) type. Firms are ex-ante identical and employ workers in entry-level and supervisory positions. Search frictions allow firms to extract rents from their supervisory positions. All experienced workers are more productive in the supervisory position, but the firm is slot constrained and may employ only one supervisory worker at any given time. The firm employs two young workers at the entry-level job and may commit to promote the highest skilled of these workers to the supervisory position when they are old.[2] This commitment is part of the entry-level contract, and in equilibrium firms never offer a contract that includes this commitment to a low-signal worker.[3] These "promotion" firms are then more likely to place a talented worker in their supervisory position than those who fill the position by hiring from the frictional market. Competition for these workers causes their wages to be bid up above their marginal products, allowing the worker to extract the benefits of this strategy.

Consequently, entry-level workers at promotion firms initially earn higher wages than those at non-promotion firms. This difference is short lived; because talent is scarce, the best workers at all firms eventually move to supervisory jobs either inside or outside their original firm. The most senior employees at entry-level jobs earn identical wages, regardless of their firm's promotion strategy, because all long-tenured entry-level employees are of low quality. Therefore, in a cross-section of *entry-level* workers, we should observe a lower return to tenure at promotion firms than at non-promotion firms. These predictions are confirmed using a matched employer-employee data set from the UK.

---

[2] Fixed firm hierarchies and slot constraints for higher level jobs are common assumptions in the tournament literature. See, for example, Waldman (2003), DeVaro and Waldman (2009) and Waldman (2011).

[3] Technically, the firm offers a contract to low-signal workers, but the wage is so low that it is never accepted.

This paper expands on the already vast literature on promotions and internal labor markets.[4] Much of the literature has focused on the value of promotions as incentives within the firm. In this tournament literature, worker effort is costly and hidden, and firms optimally design a pay hierarchy where promotions act as prizes to the best-performing worker. Typically, these papers ignore the option of the firm to hire externally.[5] One exception to this is a series of papers by Chan (1996, 2006). He shows that allowing firms to hire from outside causes the number of players in a promotion tournament to expand dramatically, lowering the incentives of the internal employees. The solution is to use lower requirements when evaluating, and thus give preference to, the internal employees for promotion. He finds empirical support using personnel records from a U.S. financial company. Tsoulouhas, Knoeber, and Agrawal (2007) show that firms should set rules that favor internal candidates unless outside candidates are of significantly higher innate ability. Waldman (2003) analyzes the commitment problem where the *ex post* optimal hiring rule over-hires outsiders, thus providing inefficient incentives relative to the *ex ante* optimal rule.

More closely related to this paper are the learning and job assignment models of promotion. The canonical models by Gibbons and Waldman (1999, 2006) were designed to explain a wide array of empirical regularities in analyses of firm personnel records.[6] In these models, there is initial uncertainty about the quality of a worker when he enters the firm. As this uncertainty is resolved, the worker is assigned to a more efficiently matched job within the firm. High ability workers will be promoted to jobs that have high marginal returns to ability, while low ability workers will remain in jobs that have limited returns to ability.[7]

---

[4] For good recent surveys of the promotion literature, see Lazear and Oyer (2009) and Waldman (2007)

[5] There has been much recent work on market-based tournaments where "prizes" are determined by the outside offers of other firms. Being promoted acts as a signal to outside firms and increases their offers for the worker. See, for example, Zabojnik and Bernhardt (2001), Ghosh and Waldman (2009), Waldman (2011), and Zabojnik (2012). For the classic promotions as signals model in a non-tournament setting, see Waldman (1984) and DeVaro and Waldman (2009).

[6] The most famous of these studies are Baker, Gibbs, and Holmstrom (1994a, 1994b). Other prominent examples include Medoff and Abraham (1980), Gibbs (1995), and Lazear (2000).

[7] Another key piece of these models is the development of human capital over time. While I ignore this in my analysis, instead constraining supervisory jobs to only be manned by older workers, I could easily modify my model to include this, by stipulating that a human capital gain with experience makes workers who were previously more productive at the entry-level job now more productive at the supervisory job.

Unlike my model, these models typically treat the firm in a vacuum, ignoring turnover and the decision of whether to consider outside workers for positions within the firm.[8]

The small literature on differences in promotion practice has thus far focused on exogenous firm heterogeneity. Using survey data from Spanish industrial plants, Bayo-Moriones and Ortin-Angel (2005) find that establishments that are likely to rely on specific capital and that face higher information asymmetries are more likely to promote internally. They find little evidence that promotions are used more frequently in establishments where providing individual incentives appears more important. DeVaro and Morita (forthcoming) model how exogenous firm productivity differences can drive the aspects of organizational structure, including promotion. They find empirical support for their theory using a questionnaire of managers included in the data set I use.

However promotion practices differ even among observationally similar firms.[9] My approach using homogenous firms provides a way to explain this variation. While previous literature has suggested that homogenous firm models may not be able explain differences in internal labor practices, my model shows that heterogeneity is not always necessary.[10] Instead market forces cause otherwise identical firms to differ in their promotion practice. Rather than deriving predictions based on firm characteristics, my model instead gives predictions based on equilibrium wage correlations.

The rest of the paper is outlined as follows. In section 2, I formally derive my model of promotion and internal labor market creation. Section 3 describes the data. In section 4, I test the predictions of my model. Section 5 concludes.

---

[8] One exception is Ghosh (2007) who derives a model where turnover is driven by match quality and firm-specific capital acquisition.

[9] For the subset of establishments of which I observe productivity measures, only 28% of the variation in promotion practice can be explained by 2-digit industry code, productivity, and the set of establishment characteristics listed in Table 2.

[10] See, for example, the discussion of hiring practices in Oyer and Schaefer (2011).

2 General Model

2.1 Primitives

I use a simple overlapping generations framework. A continuum of heterogeneous workers of Lebesgue measure 2 are born each period. Each cohort is identical. Workers live two periods and have an unknown (to both the worker and the market) type that determines their productivity. This type is either good (G) or bad (B), and, conditional on their type and job, workers are equally productive at all firms. A good worker has base productivity θ_G, while a bad worker has base productivity θ_B, where $\theta_G > \theta_B > 0$. While individual workers' types are initially unknown, the market and worker observe a signal of their quality. A worker with a high signal (H) has probability $p_H < 1$ of being a good worker, while a worker with a low signal (L) has probability $p_L < p_H$ of being a good worker. A fraction $g > 0$ of the workers have high signals, and this fraction is known *ex ante* by all firms and workers. During the second period of their life span, the true type of all workers is revealed.

Firms are infinitely lived and *ex ante* identical, and have two types of jobs: entry-level and supervisory. At the entry-level job, workers produce θ_i $(i \in G, B)$, while at the supervisory job workers produce $\lambda\theta_i$ $(\lambda > 1)$. Thus, all workers are more productive at the supervisory job than at the entry-level job but good workers more so than bad workers. Workers must first work at the entry-level job to learn the supervisory job, so only old experienced workers may be hired as supervisors. Firms are of a fixed-hierarchy; they can employ no more than one supervisory worker and no more than two young entry-level workers at any given time. For convenience, the latter restriction is relaxed for old workers. The firm can hire as many old workers for entry-level jobs as it chooses.[11] The set of potential entrant firms is of large measure greater than 1.

Each firm must additionally purchase a factory for its supervisor to produce. A continuum of identical factories of measure 1 are born each period and fully depreciate at the end of the period. Factories cost r, which is endogenously determined by the market.[12]

---

[11] Constraining the number of old entry-level workers a firm can hire would not impact the results.

Firms offer a series of contracts to attract entry-level employees. These contracts include the firm's hiring policy for its supervisory vacancy in the next period. A contract either commits the firm to fill the supervisory job the following period with its highest ability entry-level hire, or to hire the supervisor from the market.[13] Ties at firms that hire internally are broken by coin-flip. Frictions cause external hiring from the supervisory market to be done randomly. First, all good experienced workers who were not promoted internally, including workers from non-promotion firms who turned out to be good, are matched with random supervisory vacancies. Any vacancies that remain after this process are matched with random experienced bad workers.[14]

Each contract is a signal-, coworker-, and promotion policy-conditional wage offer. Workers observe all offered contracts and accept the contract which maximizes their expected lifetime utility. The firm's type is thus determined by which of its offered contracts are accepted. Firms whose accepted contracts include promotion promises will be referred to as promotion types, while firms whose accepted contracts do not will be non-promotion firms.

Note that supervisory wages are not determined by contract. Because of the friction, firms and supervisors Nash bargain over wages after they are matched. If no agreement is reached, the vacancy goes unfilled and the worker enters the market for old entry-level jobs.

Firms and workers are risk neutral and do not discount the future. Multi-period employment contracts are not permissible. Unfilled job vacancies generate 0 for the firm, and unemployed workers and unutilized factories earn 0, as well.

---

[12] The inclusion of factories is simply a device to allow for a free-entry equilibrium with zero profits. It can be viewed as analogous to an endogenous sunk cost of capital acquisition for vacancy creation in a search model context. Alternatively, I could allow firms to make profits without altering my results.

[13] This does not allow for a contract in which the firm commits to hire from the market only if the outside candidates are expected to be of higher quality than the internal candidates. This assumption is done for simplicity. Allowing for this option would create three new equilibrium wages to determine, while only serving to increase the value to the firm of options without a full promotion commitment. The main results of the model should still hold in this scenario.

[14] These frictions allow firms to extract rents from supervisory positions. In the absence of such rents, competitive wages eliminate any benefits that could be gained by manipulating employee composition to improve the future quality of supervisory jobs. Frictions for supervisory jobs are consistent with empirical work by Burdett and Cunningham (1998) which suggests that vacancies for positions with higher requirements take longer to fill. A similar assumption was also made by Waldman (2003) to model incomplete information on workers outside the firm.

2.2 Equilibrium

First, I will define equilibrium in this economy. Essentially, this concept of equilibrium is that contracts are accepted optimally, and that the firm-entry market clears.

Definition 1 *Equilibrium is set of period-specific contract outcomes* Δ *and a vector of period-specific factory costs* r *such that*

1. *Participation: The set of contracts in* Δ *that generate negative expected profits for firms or negative expected utility for workers is of measure* 0
2. *Contract Optimality: There does not exist a contract outside of* Δ *that would increase both the expected lifetime profits of a positive-measure set of firms and the expected lifetime utility of a positive-measure set of workers*
3. *Factory Market Clearing: The excess supply and excess demand for factories is non-positive in all periods, conditional on* Δ

The contract optimality condition ensures both utility maximization by the workers conditional on the set of offered contracts, and profit maximization by firms with respect to the series of contracts they choose to offer. The prevailing wage is a market clearing wage, and given that both the entry-level market and the market for supervisory capital is competitive, Lemma 2 follows.

Lemma 2 *In any equilibrium, firms make zero expected lifetime profits.*[15]

While many of the characteristics of equilibria depend on the parameters, three properties of all equilibria are immediately apparent. First, since there are no limits on the number of old workers a firm can hire for entry-level jobs, and since there is no future possibility of promotion that would cause the value of old workers to differ in entry-level jobs across firms, old workers must be paid their marginal product at entry-level jobs regardless of firm type. One can thus ignore old entry-level job workers when looking at the promotion and hiring strategies of the firms.[16]

---

[15] Proofs of all results can be found in the appendix.

Lemma 3 states a second property of all equilibria. In each period all factories are consumed. This follows from Lemma 2. Any r that is greater than 0 would generate a supply 1 of factories, and a $r = 0$ would generate positive profits to firms. This defines the set of firms active in the economy to be of measure 1 in each period.[17]

Lemma 3 *In any equilibrium, all factories are consumed in every period.*

Another property of all equilibria is stated in Lemma 4. Unemployment does not exist in this economy essentially because employing a worker always creates a surplus.

Lemma 4 *In any equilibrium, there is no unemployment in any period.*

Lemma 5 characterizes a final property of all equilibria. Entry-level workers at non-promotion firms are paid their marginal product because there is large set of potential employers for entry-level positions.

Lemma 5 *In any equilibrium, the wage of any entry-level worker employed at a non-promotion firm is equal to his base productivity.*

In order to help characterize the equilibrium I will be looking at, I will introduce the following two definitions.

Definition 6 *A steady state equilibrium is an equilibrium in which the set of firms whose accepted contracts vary over time is of measure zero.*

---

[16] This also defines the outside option in the supervisory-job wage-bargaining game as their productivity at an entry-level position. See the appendix proof of Proposition 9 for more details on equilibrium wages.

[17] Technically, this defines the set of firms with supervisors that are active in each period. Firms that do not have factories offer non-promotion contracts. Since workers are indifferent between a non-promotion contract at a firm with a supervisor and one without, I assume these contracts are never accepted in equilibrium.

As is standard in these types of dynamic models, I will be looking only for steady state equilibria, for simplicity. This will allow me to consider whether my equilibrium conditions hold in just one period of the model.

Definition 7 *A trivial equilibrium is an equilibrium in which, in some period, the set of firms that promote or the set of firms that do not promote is of measure zero.*

While trivial equilibria exist, they are not of empirical interest. In my data set I observe heterogeneity in promotion practices, so I will not consider any equilibria which cannot replicate this fact.

Definition 8 *A talent-scarce equilibrium is an equilibrium in which the set of bad workers who are employed in supervisory jobs at non-promotion firms has positive measure in every period.*

Talent scarcity is controlled by two parameters. The first is the total number of good workers that exist in the economy. The second, as will be seen in Proposition 9, is the number of workers with high signals, as this determines the number of non-promotion firms. If an equilibrium existed with just a small measure ε set of non-promotion firms then, no matter how low the probability that a worker with a given signal is good, the equilibrium would not be talent scarce by the law of large numbers.

In order to simplify notation, let ψ be the share of the surplus from supervisory jobs that goes to the firms during bargaining. I will define $\bar{\theta}_i^F = \psi(\lambda - 1)\theta_i$ as the firm's revenue from employing a worker of type i at a supervisory job. Likewise, I will define $\bar{\theta}_i^W = (1-\psi)(\lambda-1)\theta_i$ as the worker of type i's wages at a supervisory job above what they would earn at an entry-level job. Recall the set of high-signal workers in each period is measure g, and p_i, $i \in \{H, L\}$ represents the probability of being a good worker conditional on signal. Proposition 9 establishes an equilibrium in this economy.

Proposition 9 *A steady-state talent-scarce equilibrium in which the set of low-signal workers employed at entry-level jobs at promotion firms is of measure* 0*, and the set of high-signal* *workers employed at entry-level jobs at non-promotion firms is of measure* 0 *exists under the following conditions*

1. $(2p_Hp_L - p_H^2(1+\frac{g}{1-g}))(\bar{\theta}_G^F - \bar{\theta}_B^F) \geq (1-p_H)(1-p_L+p_H-p_L)(\frac{1}{2(1-g)(1-p_L)+g(1-p_H^2)})\bar{\theta}_B^W$
2. $(p_H^2\frac{g}{1-g} + p_L^2)(\bar{\theta}_G^F - \bar{\theta}_B^F) \geq (1-p_L)^2(\frac{1}{2(1-g)(1-p_L)+g(1-p_H^2)})\bar{\theta}_B^W$
3. $(2(p_H-p_L) - p_H^2(1+\frac{g}{1-g}))(\bar{\theta}_G^F - \bar{\theta}_B^F) \geq -(1-p_H)^2(\frac{1}{2(1-g)(1-p_L)+g(1-p_H^2)})\bar{\theta}_B^W$
4. $2p_L + p_H^2\frac{g}{1-g} < 1$

Proposition 9 establishes the equilibrium I will be focusing on. The first two conditions ensure that the promotion contract attracts the optimal worker composition. These essentially require that good workers are sufficiently more productive at the supervisory job than bad workers. If this is not the case, it may be optimal for the firm to offer a contract that would entice low-signal workers, who are willing to accept very low wages for an increased opportunity at getting a supervisory job when they are revealed as bad. The first condition also requires p_L to not be too low, otherwise a high-signal worker would be willing to take a substantial pay cut in order to work with a low-signal worker, which could increase the profits of the promotion firm. The third condition establishes that the high-signal workers prefer to work at promotion jobs rather than non-promotion jobs. This requires that the promotion strategy is sufficiently profitable relative to a non-promotion strategy because the future profits of the promotion firm are passed back to the entry-level workers when they are young. This condition is automatically satisfied when the promotion strategy yields better candidates for supervisory jobs than the non-promotion strategy, which is the case I will be studying for my testable implications. The final condition is the requirement that not all non-promotion firms hire good supervisory workers, and thus that the equilibrium is talent scarce.

The proposition describes a perfect sorting equilibrium. In this equilibrium, the number of promotion firms is perfectly determined by the number of high-signal workers and all high-signal workers are employed at promotion firms. As I will describe in more depth in the next subsection, as long as the signals are sufficiently different and the high-signal workers are sufficiently scarce, the motivation of the promotion firms is to get workers who are likely to be good in house at the lower ranks of the career ladder. The use of internal promotion allows them to bypass the friction involved in hiring from the market and get a better quality worker in their supervisory job in expectation.

Corollary 10 *For any set of parameters $p_H, p_L, \psi, \lambda, \theta_B$, and g, there exists some θ_G^*, such that for all $\theta_G \geq \theta_G^*$, there exists a talent-scarce equilibrium in which the set of low-signal workers employed at entry-level jobs at promotion firms is of measure zero, and the set of high-signal workers employed at entry-level jobs at non-promotion firms is of measure zero provided*

1. $2p_H p_L - p_H^2(1 + \frac{g}{1-g}) > 0$
2. $2(p_H - p_L) - p_H^2(1 + \frac{g}{1-g}) > 0$
3. $2p_L + p_H^2 \frac{g}{1-g} < 1$

While this equilibrium does not exist for all parameters, the corollary shows that this equilibrium can exist for a large range of parameters. There are several reasons why focusing on this equilibrium is appropriate. First, the trivial equilibrium clearly does not match the data. Equilibria that are not talent-scarce can exist, but these equilibria can never be the solution to a social planner's problem. On the other hand, for any social welfare function there is a set of transfers for which the equilibrium allocation in Proposition 9 will solve the social planner's problem.[18]

Further, talent scarcity is a more accurate reflection of the labor market. Discussions with employers on hiring strategies often focus on how to find and retain talent at the firm. Although other types of talent scarce equilibria exist, experimentation has suggested that these equilibria are generally unstable. These equilibria also require complicated coordination strategies by the workers which balance the number of high-signal entry-level workers at non-promotion firms, whereas in the equilibrium in Proposition 9 workers strictly prefer the contract they accept over all other types offered.[19] For these reasons, I will assume going forward that the economy is at an equilibrium as described in Proposition 9.

---

[18] For a more in-depth discussion along with proofs of these results, see Appendix B (online)

2.3 Testable Predictions

My model seeks to explain the heterogeneity in promotion practices that is observed in the data as a mechanism that is used to attract talent, and one that is limited by the talent in the labor market. Condition 11 states the parameter values under which this is the case.

Condition 11 $p_H(2-p_H) > (\frac{g}{1-g}p_H^2 + 2p_L)$

This condition ensures that the *ex ante* probability of drawing a good worker from two high-signal workers is greater than the probability of drawing a good worker from the market. Thus if Condition 11 is satisfied the promotion strategy yields better supervisory workers on average than the non-promotion strategy. There are a couple reasons to expect this condition holds in the data. First, the condition is a sufficient condition to rule out trivial equilibria in which no firms promote. Since I do not observe any industries which have homogenous promotion practices, it is likely this condition is satisfied. Second, when the expected quality of the candidate hired from the market, which is represented by the right hand side of the inequality, is high, firms may find it very profitable to renege on their promotion promise. While I do not model the commitment problem, the temptation to cheat should be highest when the condition is not satisfied.

Proposition 12 analyzes the observed wage differences among young entry-level workers. Young entry-level workers are paid more at promotion firms.

Proposition 12 *Under Condition 11, young entry-level workers at promotion firms are paid more on average than young entry-level workers at non-promotion firms*[20]

---

[19] One other stable pooling equilibrium can exist, and has similar properties to the one in Proposition 9. I discuss this equilibrium in more depth in Appendix B (online)

[20] For Propositions 12 and 14, Condition 11 is a sufficient condition with a clear interpretation. The necessary condition is $(p_H - p_L)(\theta_G - \theta_B) + \frac{1}{2}(p_H(2-p_H) - (\frac{g}{1-g}p_H^2 + 2p_L)(\bar{\theta}_G^F - \bar{\theta}_B^F)) > 0$ which takes into account that wages may still be higher at promotion firms due only to entry-level job productivity differences.

This effect is partially due to the sorting that we observe in equilibrium. Promotion firms hire exclusively high-signal workers for their entry-level jobs in equilibrium and non-promotion firms hire exclusively low-signal workers, so we would expect to see entry-level workers earn more at promotion firms. However, this is also due to labor market competition over talent. Since the promotion strategy yields better supervisory workers, firms bid up the right to hire the workers necessary to employ it. The future supervisory revenues are passed back to the entry-level workers at promotion firms in the form of higher wages when young. Workers at promotion firms are actually paid a *premium* over what they could earn in the non-promotion sector.

However, as shown in Proposition 13, this wage difference disappears for old entry-level workers. Old entry-level workers receive identical wages across firm types.

Proposition 13 *There will be no differences in average wages for old entry-level workers between promotion and non-promotion firms.*

This is for two reasons: first, since the market for old entry-level workers is perfectly competitive, workers are paid their marginal product. Second, because talent is scarce and old worker types are known, the only workers who we will observe at entry-level jobs will be bad workers. Good workers leave their entry-level jobs and take supervisory jobs, either through internal promotion or by moving to a different firm. Therefore, the average wage at entry-level jobs in the second period will not depend on the firm's type.

Combining the previous two propositions, Proposition 14 tells how the wages of entry-level workers evolve in the two different types of firms. In a cross-section of entry-level workers, the return to tenure should be higher at non-promotion firms.

Proposition 14 *Under Condition 11, the observed return to tenure of entry-level workers is higher at non-promotion firms.*

This effect is due mainly to the change in composition over time at entry-level jobs. At a promotion firm, initially the average worker's ability is fairly high, since only high-signal workers are employed. However, the fact that they are still working at the entry-level job in the second period indicates that, despite their high signal, they were revealed to be a bad worker. All low-signal workers who remain at entry-level jobs are also bad workers. But since they started with a lower expected productivity, this composition change will affect their observed average wages less dramatically. Moreover, experienced entry-level workers do not receive the wage premium paid to young workers at promotion firms because old workers provide no future profits to the firm.

Finally, since the promotion strategy yields better supervisory job workers, we should expect to see supervisory workers at promotion firms earn higher wages on average than supervisory workers at non-promotion firms.

Proposition 15 *Under Condition 11, supervisory job workers are paid more on average at promotion firms than at non-promotion firms. However, conditional on type there are no differences in wages.*

This is due entirely to sorting. Since wages at high jobs are determined by Nash bargaining and firm-type does not affect either party's threat points, if one could control for worker type (good or bad) there would be no wage difference between supervisory job wages at promotion and non-promotion firms.

To summarize these main predictions and translate them into my data, in a cross-section of workers I should observe the following:

1. Workers in entry-level jobs at promotion establishments should have higher wages conditional on promotion-policy specific tenure.

2. The observed return to tenure in a sample of entry-level job workers should be higher at non-promotion establishments than promotion establishments.

3. There should be no differences associated with the promotion policy of the employing establishment in the wages of entry-level job workers with high seniority.

4. Supervisory job workers at promotion establishments should have higher observed wages. However, this would not hold if I were able to condition on ability.

3 Data

The data come from the Worker Employment Relations Survey (WERS) 2004, a matched employer-employee data set from the UK. The survey was conducted by the Department of Trade and Industry, the Economic and Social Research Council, the Advisory, Conciliation and Arbitration Service, and the Policy Studies Institute. It is the fifth such survey in this series, which aims to study conditions in the workplace throughout Great Britain.

The survey of each establishment took place in four parts. First, the Management Questionnaire (MQ) was administered in a roughly two-hour face-to-face interview with a management representative. The Employee Representative Questionnaire (ERQ) was conducted in person with both a union and non-union employee representative, when available. This interview typically lasted 45 minutes. The Supplemental Employee Questionnaire (SEQ) was distributed to 25 employees at the firm to gather basic information on the establishment's employees. This questionnaire was self-administered. In establishments with fewer than 25 employees, all of the establishment's employees were surveyed. Finally, a Financial Questionnaire (FQ) was completed by the establishment's financial manager on the establishment's recent financial performance. For this paper, data from the MQ and SEQ are linked to form a sample of employees matched with characteristics of their employing establishments, as reported by management.

One shortcoming of the SEQ is that many of the variables are categorical. This is particularly problematic for the wage and income data. The survey asks the workers for their weekly income, hourly wage, and number of hours they work per week. Workers are given the choice of 14 weekly income categories, but only 4 wage categories. The wage categories are also very unbalanced. The third highest (second lowest) category is for hourly wages between $\pounds 4.50 - \pounds 5.00$, while the second highest category has a range of £5.00-£15.00.

I impute hourly wages by dividing the midpoint of the weekly income categories by the continuously measured weekly hours variable. For the unbounded category, which is weekly income £871 or more, I use the value 950. I include an indicator in any regression that involves wage for whether the wage was imputed from a top-coded income.[21] This requires that the measurement error due to the categorical transformation is orthogonal to the estimated value, so that the mean is roughly equal to the midpoint of each category, and that hours worked is orthogonal to any deviations in income from the midpoint. While the assumption is strong, it is commonly made.[22]

The data consist of 22,451 workers surveyed from 2,295 unique establishments. Included in these data are 4-digit occupation and industry codes, as well as a variety of questions in the MQ that are linked to the specific 1-digit occupation codes within that establishment. In order to focus on establishment practices, rather than the results of collective bargaining, I drop all union members from the data set.[23] I also eliminate all workers who work less than 30 hours per week or whose imputed wages are less than £1/hour. I drop all workers who do not have a valid response for ethnicity, gender, age, education, occupation, and supervisor status, and anyone employed at an establishment that does not have a valid industry code. After modifying my sample in this way, I am left with 9,348 unique workers from 1,533 different establishments.

I take my promotion policy measure from the MQ. The question asks "Which of these statements best describes your approach to filling vacancies at this workplace?" Managers then chose whether internal candidates or external candidates were given preferences for vacancies, or whether both types of candidates were given equal opportunity. Roughly 93% of the establishments weighted by employment in the survey reported that they either gave preference to internal candidates or gave no preference to either type. Throughout the paper I classify establishments that report favoring internal candidates as promoting internally as in the model. This same question was previously used by DeVaro and Morita (forthcoming) to ascertain an establishment's promotion practices.

---

[21] In practice, this affects only 6.5% of my final weighted sample, and less than 1.5% of my final weighted sample of "entry-level" workers.

[22] As a robustness check, I used ordered probits with the cut points constrained to match the actual (log) income categorical cutoffs in the survey. I exploit the fact that $\ln(\text{Wage}) = \ln(\text{Weekly Income}) - \ln(\text{Hours})$, by controlling for the log of hours worked and constraining the coefficient on this variable to be equal to one. Thus all coefficients are identified and are marginal wage effects. This requires an alternative, but equally strong assumption that errors in wage are distributed log-normally. This approach yields similar results, though my tests have less power than the midpoint approach. These results are available upon request.

[23] I do this as it is the finest definition of union membership that I have, and therefore will allow me to keep the most observations. Using alternative definitions of union influence, as well as including all union members in my regressions, does not meaningfully affect the results.

The model's predictions are focused only on a subset of the jobs that are observed at a firm. In order to test these, I divide this sample into two categories. I classify "entry-level" jobs as non-managerial and non-supervisorial jobs, while managers and supervisors correspond to my model's supervisory job.[24] In Table 1, I present summary statistics for each sub-sample using the weights provided by the survey makers for the SEQ. Education in the survey is given in terms of certificates earned through the UK schooling system. In the table I reduce the seven categories given in the survey to five, though I use all seven as controls in regressions that use worker characteristics.[25] While worker characteristics are somewhat different between promotion and non-promotion establishments, the distribution of occupations is quite similar. In both sub-samples, workers at promotion establishments are more educated and earn higher wages. I can reject the equality of the distribution of education using a Kolmogorov-Smirnov test. The differences in worker occupations appears to be small. The biggest differences are among professionals, skilled trades, and personal service, which encompasses occupations such as nurses, hairdressers, and travel agents. Workers at promotion establishments are more likely to be classified as professionals while being less likely to be in the latter two, particularly at the entry-level. While the distribution of tenure among promotion supervisors does not stochastically dominate that of non-promotion supervisors, they are more likely to have 2-5 years of tenure and less likely to have less than 1 year of tenure.

---

[24] This is slightly different from DeVaro and Morita (forthcoming). They classify professionals as high level jobs, and do not have access to information on supervisors, since it is only available at the worker level. My classification is more appropriate in this setting, since it is unlikely that one would be promoted into the position of lawyer, while one could be promoted from a lawyer to a managing partner. Reclassifying professionals as supervisorial jobs yields similar results.

[25] Workers were asked to check each of the categories of education certifications they possessed. The seven categories are: no academic qualifications, GCSE grades D-G and equivalents, GCSE grades A-C and equivalents, 1 GCE 'A' level grades A-E or 1-2 SCE Higher grades A-C or AS levels, 2 or more GCE 'A' levels A-E or 3 or more SCE Higher grades A-C, First Degree, and Higher (Postgraduate) Degree. Respondents were also given the option to choose "other academic certification." I do not use that category since it is indeterminate whether that certification is higher or lower than the others chosen by the respondent.

Table 2 shows descriptive statistics at the establishment level, using weights designed to be representative of the average worker's place of employment.[26] Consistent with previous studies (DeVaro and Morita, forthcoming), promotion establishments are larger and appear to be part of firms which employ more people in the United Kingdom. This latter difference is not statistically significant. There is no difference in the age of the establishment across promotion types, suggesting that the decision to promote is not related to being part of a newer, more innovative establishment. The percentage of the establishment's employees who are trade union members is also identical across both promotion types. Promotion establishments are much more likely to offer incentive pay to at least some employees. The table lists the distribution for the largest non-managerial 1-digit occupation code at the establishment. While these statistics vary between the promotion and non-promotion establishments, there is no single type of occupation that is unrepresented in one of the promotion strategies.[27] The lowest representation is for the Personal Service category in promotion establishments, which makes up 4% of that type.

4 Empirical Results

4.1 Main Results

I first test whether entry-level job workers at promotion establishments earn higher wages initially and have a lower observed return to tenure. I estimate the following equation

$$\ln \text{Wage}_i = \beta_0 + \beta_1 X_i + B_2 * \text{Promotion} + \beta_3 * \text{Tenure} + \beta_4 * \text{Promotion} * \text{Tenure} + \epsilon_i \quad (1)$$

---

[26] The alternative set of establishment weights, which are designed for analysis of the average establishment, weight small workplaces much more highly due to their prevalence in spite of the low fraction of employment they represent.

[27] There are differences between promotion and non-promotion establishments in their 1-digit industry classification, as one would expect if each industry is a separate labor market in a separate equilibrium. However, I likewise observe both promotion and non-promotion establishments in each industry.

where X_i is a vector of controls and promotion is an indicator for whether the establishment practices internal promotion. My prediction is that $\beta_2 > 0$ and $\beta_4 < 0$. I transform the tenure categories into a continuous variable using the midpoints of the bins, and include a dummy variable for those who are in the top bin.[28] I use my sample of entry-level job workers, which consists of non-managerial and non-supervisorial workers.

Table 3 shows the results of this regression. The first column contains only the promotion dummy, tenure, and the tenure interaction with no additional controls. The coefficients on both promotion and the interaction match up with the prediction and are jointly significant at the 1% level. Non-managerial, non-supervisorial workers at establishments that promote earn on average 14% more than those at establishments that do not promote, conditional on tenure. The relevant comparison for the model, however, is for workers at identical jobs. Therefore, column 2 controls for 3-digit occupation code from the UK Standard Occupation Classification 2000 system (SOC2000) and 3-digit industry code from the UK Standard Industry 2003 system (SIC2003). Consistent with the predictions, the premium paid by promotion establishments is 6%, but these non-managerial, non-supervisorial workers have a 0.8 percentage point lower observed return to tenure and both coefficients are statistically significant. The return to tenure at promotion establishments is only slightly more than half that at non-promotion establishments.[29]

Since my predictions reflect both equilibrium sorting and a wage premium paid at promotion establishments, one would expect they would also hold when controlling for worker characteristics.[30] In column 3, I replace the industry and occupation codes with controls for education, ethnicity, gender, and age.[31] Both coefficients of interest have the right sign and are significant. Column 4 adds the industry and occupation codes back into the regression with the worker controls. The worker controls do not substantially alter the results when compared to the regression with just the industry and occupation codes, though the tenure interaction is now somewhat stronger. If these controls are reducing the sorting effect, we would expect the magnitude of the coefficients on both promotion and the interaction to decrease. This is indeed the case for the promotion coefficient. Comparing column 3 to column 1 and column 4 to column 2, the observed effect of being at a promotion firm is smaller, and significantly so in the former case at the 10% level. The coefficient on the interaction moves in the wrong direction when comparing column 1 to column 3, though this change is not significant. There is virtually no change in the interaction term when comparing 2 to 4.

---

[28] The results are robust to using both the upper or lower bounds of the bin, alternatively, to define tenure. Specifications using the tenure categories are also generally supportive of the model.

[29] It is likely that there are things outside the model that would cause differences in the return to tenure at the industry and occupation level. The results are robust to interacting tenure with both sets of codes.

[30] Technically, in my model, there is only one observable worker characteristic, the pre-market signal, and there is perfect sorting on that signal. Therefore, the effects of that characteristic and being at a promotion establishment are not separately identifiable. However, one would expect that in the data there will be some mixing of worker types across establishments, and a more robust model would allow for this prediction.

[31] The education controls include 6 dummies to represent, from lowest to highest achievement: No academic qualifications, GCSE grade D-G equivalent, GCSE grade A-C equivalent, 1 GCE A-level equivalent, 2 or more GCE A-level equivalents, Bachelor's degree equivalent, and graduate degree. The ethnicity controls are

Finally, in column 5, I add controls for the region of the establishment as well as the log number of its employees.[32] The establishment-size wage premium has been well documented and, as seen in the descriptive statistics, establishments that promote are larger on average than those who do not. One concern is that the promotion variable is simply proxying for establishment size. Though establishment size does appear to have a positive impact on wages, adding this control has little effect on the coefficients of interest. Alternatively using firm size in the UK, as well as interacting tenure with establishment size does not affect the results.

My model makes a specific prediction about the wages of the most senior entry-level job workers. The higher wages at promotion firms are due to both the higher productivity of their workers and a premium paid due to the higher expected future value they will bring the firm upon promotion. Old workers who are of high ability, however, leave the entry-level jobs due either to internal promotion or to taking a higher position at a competing firm. Further, old workers die at the end of the period; they provide no future value to the firm once the possibility of promotion is removed. Thus, there is no premium above marginal

---

16 dummies that represent British, Irish, other white, white and Black Caribbean, white and Black African, white and Asian, other mixed, Indian, Pakistani, Bangladeshi, other (South) Asian, Caribbean, African, other Black, Chinese, other ethnic group, and those who coded multiple of these.

[32] The results are robust to using levels instead of logs.

product paid to old entry-level workers at promotion firms. The wages paid to workers at entry-level jobs then should be identical at both types of firms.

In Table 4, I test this using the sample of non-managerial, non-supervisorial workers who are in the highest tenure bin, which is 10 years or higher.[33] Column 1 shows that without any controls high seniority entry-level job workers at promotion establishments earn 14 percent more on average than those at non-promotion establishments. However, this is entirely explained by differences in industry and occupation. Adding industry and occupation controls reduces the magnitude to a statistically insignificant 3%. Column 3 adds worker characteristics as controls which has little impact on the point estimate for working at a promotion establishment. This is largely to be expected, since in my model the true productivity of a worker is represented in the long run by her job, and not by her initial observable characteristics. The results are further robust to controlling for region and establishment size in column 4.

The fourth prediction of the model is that workers in supervisorial jobs should have higher observed wages at promotion establishments, which is entirely due to sorting. Note that this prediction is not conditional on tenure; promotion supervisors should also have higher tenure since in the model they worked a year at entry-level at the same firm.[34] The first column of Table 5 estimates the effect of establishment promotion policy on the wages of managers and supervisors. The coefficient on being at a promotion establishment is positive and significant. On average, managers and supervisors at establishments that promote earn 7.4 percent more than at establishments that do not promote.

Column 2 adds in 3-digit occupation and industry codes. Adding these controls reduces the magnitude of the coefficient to an insignificant .02. Controlling for worker characteristics and geographic fixed effects in column 3 further reduces the coefficient to .01. It is not entirely clear that this contradicts my model. I should not see any wage difference if I were able to condition on ability. It is possible that occupation and industry may contain information on the ability of experienced workers. Suppose, as an example, that a lumber firm consists of two establishments: a farm supervised by a foreman and a mill supervised by a technician. The marginal return to ability as a technician is higher than that of a foreman, so when a high ability supervisor is available, he is assigned to the mill. These two occupations would be classified differently under 3-digit industry and occupation codes, although both may draw from the same initial pool of workers. In fact, when I condition on broader 2-digit codes which would encompass more related activities, I find a strong positive wage premium for working at a promotion establishment.

---

[33] Results are similar when a high seniority worker is defined as 5 years of tenure or more.

[34] The effect of tenure at supervisory jobs is not separately identifiable from firm type. All supervisory job workers at promotion firms have a tenure of one, while all supervisory job workers at non-promotion firms have a tenure of zero.

Furthermore, while my promotion measure is a good measure of overall firm policy, it is only a noisy measure of whether an individual supervisor was hired internally. If promotion firms sometimes hire external candidates for reasons outside my model, these workers will be drawn from the same pool of workers as those firms that hire primarily externally. There should be no wage differences among this set of workers, which bias my results toward zero. That I see a positive coefficient on tenure when added to the regression in the fourth and fifth columns, while unsurprising, suggests that those who are promoted internally are paid more, since they trivially have more tenure.[35]

Much of the predicted wage differences between workers at promotion and non-promotion establishments is due to sorting. If one could observe the signal received by the market in my model, the entry-level jobs at promotion establishments would be entirely populated by workers with the high signal, while those jobs at non-promotion establishments would be entirely populated by workers with the low signal. Workers at supervisory jobs in promotion establishments would also all have had the high signal, since they are chosen only from the population of lower workers at such establishments. However, supervisory jobs at non-promotion establishments would be populated by both kinds of workers, since some of the workers who had high signals and were realized to be good migrate to supervisory jobs at non-promotion establishments. There is no such question, unfortunately, in the survey. Instead, I use ordered logits to look for sorting on education, an observable characteristic that is correlated with higher levels of productivity and presumably a better market signal.

---

[35] Hassink and Russo (2008) show that workers who are promoted internally are paid more and appear to be of higher quality than those hired from outside the firm, consistent with this prediction of my model.

The first column of Table 6 looks at all workers controlling only for the employer's promotion policy type. Workers at promotion establishments are significantly more likely to have higher levels of education than those at non-promotion establishments. This is true even when controlling for industry, occupation, geography and other worker characteristics. The third and fourth columns break the sorting down by job type. While both supervisory and entry-level job workers are more likely to have higher levels of education at promotion establishments, this effect is only statistically significant for the latter. To test they hypothesis that there is more sorting on education in entry-level jobs than supervisory jobs, I estimate an ordered logit that fully nests both job types. Though not statistically significant, the interaction between entry-level job status and working at a promotion establishment is positive, which is consistent with the theory.

4.2 Productivity and Incentives

One major concern is that promotion could be proxying for some sort of establishment-level productivity differences. The WERS FQ offers three different measures of establishment level productivity: profits, value added, and the capital/labor ratio. Capital/labor ratio and profitability, in particular, have been shown to be positive correlates of the wage, at least at the industry level (Dickens and Katz, 1986). Unfortunately, the response rate to the FQ was much lower than the other parts of the survey; less than half of the participating companies filled out a FQ. This severely reduces the sample size. Only 1,791 of my non-managerial non-supervisorial workers work at an establishment with a valid entry for capital/labor ratio.

Table 7 shows the results of adding these controls to the log wage regressions of non-managerial, non-supervisorial workers.[36] The results largely support the model.[37] While the coefficient on promotion is insignificant when controlling for profitability and value added, it is estimated with less precision due to the smaller sample size and is not significantly different from the estimate in column 4 of Table 3. The interaction between being at a promotion establishment and tenure is always significant and is larger, though not significantly so, than estimates seen without these controls previously. In each specification, I can reject the joint hypothesis that both coefficients of interest are equal to 0.

So far the results I have shown have been consistent with the implications of the model. They also may seem somewhat consistent with a story of promotions as incentives. In a typical promotions as tournaments model, those who do not get promoted receive lower wages as a punishment to induce higher effort in the initial period. A lower return to tenure from promotion "losers" is consistent with this. It is more difficult to explain the higher initial wages at a promotion establishment from this perspective.

One way to do so would be through a promotions as signals story. Suppose that firms have private information on the quality of their workers, and that promotion serves to signal the quality of the promoted worker to the market. Market forces then cause an increase in that individual's wages. Promotion losers will be paid less, since Bayesian updating implies they must be of lower average productivity than before. Workers at promotion firms are then paid a premium to account for risk aversion. This premium is acceptable to the firm because the incentives cause higher effort and thus higher productivity. However, the fact that long-tenured non-managerial, non-supervisorial workers appear to be paid equally across promotion strategy types is inconsistent with this approach. The workers in the non-promotion firms should be of the same average ability as the market initially viewed them. Any hiring away from this pool would have to have been done randomly and would not affect the market's beliefs. The workers at the non-promotion firms have been identified as being of lower than average skills since they lost the tournament. Since the uncertainty has been resolved, they no longer must be compensated for risk. This signals model predicts *lower* wages for long-tenured entry-level workers at promotion firms than similar workers at non-promotion firms, which is inconsistent with the results of Table 4.[38]

---

[36] There are a few theoretical problems with adding firm productivity measures to a wage regression. In a competitive market, wages are equal to marginal product. If the production function is Cobb-Douglas, log wage is a linear function of just the log of the capital/labor ratio. Value added (revenue - variable cost) divided by labor is just a linear transformation of the capital/labor ratio, as is profits less fixed costs divided by worker. In this setting, this type of regression would simply be of wage on wage. My model's setting is slightly different, in that young promotion firm workers at entry-level jobs earn wages above their marginal product in equilibrium, but this rent disappears for old workers and it is still unclear what further controls in the regression could account for. Despite these concerns, this is a regression which is frequently estimated, and can still provide some insight into the robustness of my main results.

Another possible concern is that these productivity differences are outcomes of firm hiring policy, rather than firm characteristics. Any difference in the productivity measures I use, however, would have to come from factors outside my model. Each firm in equilibrium earns equal profits (0) and employs the same amount of capital (one factory). Firms do vary in size in my model, but since I control for log establishment size and because ratios are linear in logs, these differences will not confound my results.

[37] Table 7 uses logs of these variables. The results are robust to using levels.

One additional way to address the issue of incentives is to look at the interaction between incentive pay and promotion policy. If incentive pay and promotion incentives are substitutes, and the effects I observe are only due to the implementation of a system of promotion-based pay incentives, then I should not observe a lower return to tenure at establishments that hire internally and offer incentive pay.[39] I find that my results hold for both establishments that do and do not offer incentive pay, which lends further support that these effects are due to the equilibrium sorting and hiring strategy predictions of the model.

5 Summary and Conclusion

In this paper, I developed a new model of promotions in equilibrium, which sought to explain heterogeneity in promotion practices across firms. My model is based on ex-ante identical firms committing to eventually promote one of their workers to a more valuable task before workers are hired, in order to attract talented workers into low level jobs and increase their probability of eventually obtaining a highly productive worker at the valuable task. Internal promotion allows firms to avoid the uncertainties of hiring from a frictional labor market. The viability of this strategy is limited by the scarcity of talented workers in the initial labor pool.

---

[38] This argument is based on the wage structure in Zabojnik and Bernhardt (2001), where the existence of sufficient amounts of firm specific capital prevent a market unraveling due to the asymmetric information with simultaneous wage offers. In the offer matching context of Ghosh and Waldman (2010), it is unclear how a non-promotion firm could attract any workers. In the second period, workers are paid the wage of the productivity of the lowest type of worker, and in the first period competition sends these expected rents back to the workers. Since the promotion rule is designed to maximize rents, wages at promotion firms would be strictly higher in the first period, independent of effort choice, and identical in the second, meaning that workers would strictly prefer to work at a promotion firm. A perhaps larger problem with allowing for non-promotion firms in these models is the lack of equilibrium turnover from which a non-promotion firm could hire.

[39] Recent theoretical and empirical work (Kwon, 2006, Frederiksen and Takats, 2011) has suggested that promotion incentives and other monetary bonuses are complements. If this is the case and my results are only the result of systems of promotion-bases pay incentives, then I should expect to see my results hold only on the subset of establishments that do offer incentive pay.

My model predicted that workers at entry-level jobs at firms that promote would have higher wages, but lower observed returns to tenure. This was due to two effects. First there is a sorting effect. Workers with good future prospects were initially employed only at promotion firms, but the best workers at entry-level jobs left those jobs for better opportunities, either internally or externally. Second, initial labor market competition bid up the wages of workers with good future prospects above their marginal product. Once the uncertainty has been resolved, this promotion premium disappears. There are no differences in the wages of the most senior workers in jobs that differ only in their firm's promotion policy. I confirmed these predictions in the data. Also consistent with my model, workers in supervisory jobs at promotion establishments earn higher wages and there is sorting on observables, where educated workers are more likely to be employed at promotion establishments than non-promotion establishments. These results cannot be explained by the conventional promotions as incentives models alone and are not driven by differences in establishment-level productivity.

While my model does not incorporate any incentive structure, it is compatible with the tournament literature. Firms' promotion decisions were based simply on hiring the most productive worker. One could imagine an extension which incorporated intra-firm competition for human capital in order to determine promotion. However, the decisions of the non-promotion firms must also be taken into account. If workers differed in both their expected ability and their expected return on human capital investment, similar to the job assignment literature (Gibbons and Waldman, 1999, 2006), then sorting may still be possible. This is certainly an avenue for future study.

My model also provides a potential explanation for inequality outcomes. Workers at promotion firms, who all have good observables in equilibrium, have a higher chance of earning a high-paying job even when they are revealed to not be of high ability. Applying this to the statistical discrimination literature, if blacks have poorer characteristics than whites that are observable to employers but not to the econometrician then, even conditional on other observables and ability, blacks may perform poorer than whites in the long run. This is consistent with evidence presented by Altonji and Pierret (2001).

Overall, while incentives may play an important role in the labor market within a firm, incentivizing workers to enter the firm is important when looking between firms that do and do not promote. The quality of the labor pool is a constraint on the number of firms who can profitably implement an internal labor market.

References

Altonji, Joseph G. and Charles R. Pierret. 2001. “Employer Learning and Statistical Discrimination.” *The Quarterly Journal of Economics* 116 (1): 313-350.

Baker, George, Michael Gibbs and Bengt Holmstrom. 1994. “The Internal Economics of the Firm: Evidence from Personnel Data.” *The Quarterly Journal of Economics* 109 (4): 881-919.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994. “The Wage Policy of a Firm.” *The Quarterly Journal of Economics* 109 (4): 921-955.

Bayo-Moriones, Alberto and Pedro Ortin-Angel. 2006. “Internal Promotion versus External Recruitment in Industrial Plants in Spain.” *Industrial and Labor Relations Review* 59 (3): 451-470.

Burdett, Kenneth and Elizabeth J. Cunningham. 1998. “Toward a Theory of Vacancies.” *Journal of Labor Economics* 16 (3): 445-478.

Chan, William. 1996. “External Recruitment versus Internal Promotion.” *Journal of Labor Economics* 14 (4): 555-570.

Chan, William. 2006. “External Recruitment and Intrafirm Mobility.” *Economic Inquiry* 44 (1): 169-184

Devaro, Jed and Hodaka Morita. (forthcoming.) “Internal Promotion and External Recruitment: A Theoretical and Empirical Analysis.” *Journal of Labor Economics.*

Devaro, Jed and Michael Waldman. 2009. “The Signaling Role of Promotions: Further Theory and Evidence.” Cornell University ILR School Working Paper 8.

Dickens, William T. and Lawrence F. Katz. 1986. “Interindustry Wage Differences and Industry Characteristics.” NBER Working Paper No. 2014

Frederiksen, Anders, and Elod Takats. 2011. “Promotions, Dismissals, and Employee Selection: Theory and Evidence.” *The Journal of Law, Economics, and Organization* 27 (1): 159-179.

Ghosh, Suman. 2007. “Job mobility and careers in firms.” *Labour Economics* 14 (3): 603-621.

Ghosh, Suman and Michael Waldman. 2010. “Standard promotion practices versus up or out contracts.” *The RAND Journal of Economics* 41 (2): 301-325.

Gibbons, Robert and Michael Waldman. 1999. “A Theory of Wage and Promotion Dynamics Inside Firms.” *The Quarterly Journal of Economics* 114 (4): 1321-1358.

Gibbons, Robert and Michael Waldman. 2006. “Enriching a Theory of Wage and Promotion Dynamics Inside Firms.” *Journal of Labor Economics* 24 (1): 59-107.

Gibbs, Michael. 1995. “Incentive compensation in a corporate hierarchy.” *Journal of Accounting and Economics* 19 (2-3): 247-277.

Hassink, Wolter and Giovanni Russo. 2008. “Wage differences between internal and external candidates.” *International Journal of Manpower* 29 (8): 715-730.

Kwon, Illoong. 2006. “Incentives, wages, and promotions: theory and evidence.” *The RAND Journal of Economics* 37 (1): 100-120.

Lazear, Edward P. “Performance Pay and Productivity.” *The American Economic Review* 90 (5): 1346-1361.

Lazear, Edward P. and Paul Oyer. 2009. “Personnel Economics.” Unpublished manuscript, Stanford Graduate School of Business.

Lazear, Edward P. and Sherwin Rosen. 1981. “Rank-Order Tournaments as Optimum Labor Contracts.” *The Journal of Political Economy* 89 (5): 841-864.

Li, Hao and Sherwin Rosen. 1998. “Unraveling in Matching Markets.” The American Economic Review 88 (3): 371-387.

Medoff, James L. and Katharine G. Abraham. 1980. “Experience, Performance, and Earnings.” *The Quarterly Journal of Economics* 95 (4): 703-736.

Oyer, Paul and Scott Schaefer. 2011. “Personell Economics: Hiring and Incentives.” in *Handbook of Labor Economics*, eds. Ashenfelter, Orley and David Card. Elsevier: Amsterdam.

Tsoulouhas, Theofanis, Charles R. Knoeber and Anup Agrawal. 2007. “Contests to become CEO: incentives, selection and handicaps.” *Economic Theory* 30 (2): 195-221.

Waldman, Michael. 1984. “Job Assignments, Signaling, and Efficiency.” *The Rand Journal of Economics* 15 (2): 255-267.

Waldman, Michael. 2003. “Ex Ante Versus Ex Post Optimal Promotion Rules: the Case of Internal Promotion.” *Economic Inquiry* 41 (1): 27-41.

Waldman, Michael. 2007. “Theory and evidence in internal labor markets.” MPRA Paper No. 5113

Waldman, Michael. 2011. “Classic Promotion Tournaments versus Market Based Tournaments.” Unpublished manuscript, Cornell University.

Department of Trade and Industry. Employment Markets Analysis and Research et al. , *Workplace Employee Relations Survey, 2004: Cross-Section Survey, 2004 and Panel Survey, 1998-2004; Wave 2* [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], January 2008. SN: 5294.

Zabojnik, Jan. 2012. “Promotion tournaments in market equilibrium.” *Economic Theory* 51 (1): 213-240.

Zabojnik, Jan and Dan Bernhardt. 2001. “Corporate Tournaments, Human Capital Acquisition, and the Firm-Size Wage Relation.” *Review of Economic Studies* 68 (3): 693-716.

A Proofs of Main Results

A.1 Proof of Lemma 2

Proof. Suppose in some period there was a contract in Δ that would generate positive profit given r_t . Then all firms demand a factory at price r_t. Since the supply of factories is fixed at 1, there is excess demand for factories. This violates Factory Market Clearing. ■

A.2 Proof of Lemma 3

Proof. Suppose there was a period t in which some factories were not consumed. Then it must be that r_t is greater than the expected benefit of hiring a supervisory worker from the market. Otherwise, firms could make positive lifetime profits by entering in t only and hiring a supervisory worker from the market, which is not possible via Lemma 2. Since $\theta_B > 0$, all workers have positive production, and since supervisory wages are determined by Nash bargaining, the expected benefit to the firm of hiring a supervisory worker is positive, this means there is excess supply of factories at the price r_t. This violates Factory Market Clearing. ■

A.3 Proof of Lemma 4

Proof. Suppose there was unemployment in some period among old workers. Then it must be that the wage for old entry-level workers is 0, otherwise there is an excess supply of old workers. However, since $\theta_B > 0$, this implies that firms can make positive profit in that period by hiring only old entry-level workers. This violates Lemma 2.

Suppose there was unemployment in some period among young, low-signal workers. If the wage of low-signal workers at non-promotion firms in that period gives positive utility, then there is some contract outside of Δ with a lower wage that would increase the profits of a firm and the utility of an unemployed worker.

Suppose the wage gives 0 utility to low-signal workers in entry-level jobs at non-promotion firms in some period, and some low-signal workers are unemployed. Since low-signal workers generate positive production in expectation for the firm, there is some contract outside of Δ that offers a slightly higher wage, would generate higher utility for the unemployed workers and profits for firms that do not employ entry-level workers (non-entrant firms). This violates Optimal Contracts. An analogous argument holds for high-signal workers. ■

A.4 Proof of Lemma 5

Proof. Suppose in some period there is a contract in Δ that has no promotion promise and a wage above the worker's base productivity. The the firm takes a loss on the contract which violates Participation. Suppose in some period there is a contract in Δ that has no promotion promise and a wage below the worker's productivity. Then there exists a higher wage contract which would increase the utility of the worker and the profits of a (non-entrant) firm which does not employ entry-level workers. This violates Optimal Contracts. ■

A.5 Proof of Proposition 9

Proof. By Lemma 4 there is no unemployment, so the set of promotion firms must be of measure g. By Lemma 3, all factories are consumed, so there then must be $1 - g$ non-promotion firms who each employ 2 low-signal workers at entry-level jobs. Now consider the expected quality of hiring a supervisory worker. Given this allocation of workers and firms, there is excess demand for good workers at supervisory jobs at non-promotion firms because

$$1-g > gp_H^2 + 2(1-g)p_L \tag{2}$$

where the first term on the right-hand side is the measure of the set of good workers who do not get supervisory jobs at promotion firms and the second term is the measure of the set of good workers who were previously employed at entry-level jobs in the non-promotion firms. Rearranging these terms, the probability of hiring a good worker for a supervisory job at a non-promotion firm is

$$\frac{g}{1-g}p_H^2 + 2p_L \tag{3}$$

Condition 4 states that this must be less than 1, and ensures talent scarcity.

Given this, and that the wage of entry-level workers at non-promotion firms is equal to marginal product (Lemma 5), Lemma 2 determines the price of factories r. The expected profits for a non-promotion firm are

$$\pi^{np} = \bar{\theta}_B^F + (\frac{g}{1-g}p_H^2 + 2p_L)(\bar{\theta}_G^F - \bar{\theta}_B^F) - r \tag{4}$$

Setting $\pi^p = 0$ and rearranging terms determines r.

Finally, the wage for the high-signal workers at entry-level jobs at promotion firm must induce zero profits by Lemma 2. The expected output from entry-level workers is $2(p_H\theta_G + (1-p_H)\theta_B)$, and the expected probability of getting a good worker for the supervisory job is $p_H(2-p_H)$. The promotion firm must pay $\bar{\theta}_B^F + (\frac{g}{1-g}p_H^2 + 2p_L)(\bar{\theta}_G^F - \bar{\theta}_B^F)$ to purchase a factory. Denote w^{HH} as the wages paid by promotion firms to high-signal workers with a high-signal coworker. The promotion firm's expected profits from hiring two high-signal workers are

$$\pi^p = 2(p_H\theta_G + (1-p_H)\theta_B) + (p_H(2-p_H) - (\frac{g}{1-g}p_H^2 + 2p_L)(\bar{\theta}_G^F - \bar{\theta}_B^F)) - 2w^{HH} \tag{5}$$

Setting $\pi^p = 0$ rearranging terms determines the equilibrium wage w^{HH}.

In order for this to be an equilibrium, it must be that there are no other contracts that would be strictly preferred by both a worker and a firm. I will show this by finding the lowest wage that would be accepted by a worker at each job, and finding under which conditions the firm would not be willing to offer that wage.

There are two other entry-level worker combination contracts offered by the promotion firm. The first mixes one high-signal worker with one low-signal worker. The second employs two low-signal workers. A worker will always get a supervisory job when she is revealed to be good, due to talent scarcity, but her probability of getting a supervisory job when she is revealed to be bad varies with the job. Since ties are broken by coin flip, workers at a promotion firm with one low- and one high-signal entry-level job worker have probability $\frac{1}{2}(1-p_L)(1-p_H)$ of getting a promotion internally as a bad worker. That probability is $\frac{1}{2}(1-p_L)^2$ and $\frac{1}{2}(1-p_H)^2$ for two high-signal and two-low signal entry-level job workers, respectively. Workers also have a chance of getting a supervisory job on the market when they are bad because of talent scarcity. Denote this probability as χ.

In order to accept an alternative entry-level contract, a worker must receive a wage differential that is equal to the change in her probability of obtaining a supervisory job when old. Let $\tilde{w}^{ij}$ be the lowest entry-level wage a worker with signal i would accept for a promotion job with co-worker j, and $\tilde{w}_i$ be the lowest entry-level wage a worker with signal i would accept at a non-promotion job. These wages are

$$\tilde{w}^{HL} = w^{HH} - \frac{1}{2}(1-p_L)(p_H - p_L)(1-\chi)\bar{\theta}_B^W \tag{6}$$

$$\tilde{w}^{LH} = w_L - \frac{1}{2}(1-p_H)(1-p_L)(1-\chi)\bar{\theta}_B^W \tag{7}$$

$$\tilde{w}^{LL} = w_L - \frac{1}{2}(1-p_L)^2(1-\chi)\bar{\theta}_B^W \tag{8}$$

$$\tilde{w}_H = w^{HH} + \frac{1}{2}(1-p_H)^2(1-\chi)\bar{\theta}_B^W \tag{9}$$

These wages are simply the equilibrium wages from the contract the worker accepts less the gain in supervisorial probability he receives from switching contracts times the added wage of a bad supervisor.

A promotion firm that hires one high- and one low-signal worker has probability $(p_H + p_L - p_H p_L)$ of hiring a good worker for its supervisory job. Thus, the firm's expected profits from this strategy, π_{HL}^p, are

$$\pi_{HL}^p = (p_H+p_L)\theta_G+(2-p_H-p_L)\theta_B+(p_H(1-p_L)-\frac{g}{1-g}p_H^2-p_L)(\bar{\theta}_G^F-\bar{\theta}_B^F)-(w^{HL}+w^{LH}) \tag{10}$$

For this to be an equilibrium, it must be that the profits at the lowest wages workers are willing to accept are less than zero for a promotion firm to hire one high-signal and one low-signal worker. Substituting for $\tilde{w}^{HL}$ and $\tilde{w}^{LH}$, this is true if and only if

$$\frac{1}{2}(1-p_H)(p_H-p_L+1-p_L)(1-\chi)\bar{\theta}_B^W - (p_H p_L + \frac{1}{2}(1+\frac{g}{1-g})p_H^2)(\bar{\theta}_G^F - \bar{\theta}_B^F) \leq 0 \tag{11}$$

The hereto undefined parameter, χ, represents the probability of finding a supervisory job in the market conditional on being a bad worker. There is an oversupply of bad workers for good jobs because

$$g(1-p_H^2) + 2(1-p_L)(1-g) > (1-g)(1-\frac{g}{1-g}p_H^2 - 2p_L) \tag{12}$$

where the first term on the left-hand side is the measure of the set of bad workers at promotion firms who are not promoted, the second term on the left-hand side is the measure of the set of bad workers that were previously at entry-level jobs at non-promotion firms, and the right-hand side is the measure of the set of supervisory vacancies that are not filled by good workers at non-promotion firms. Re-arranging,

$$\chi = \frac{(1-g)(1-\frac{g}{1-g}p_H^2 - 2p_L)}{g(1-p_H^2) + 2(1-p_L)(1-g)} \tag{13}$$

Substituting in for χ the above inequality is condition 1 in the proposition.

Focusing now on the second alternative contract offered by a promotion firm, when it hires two low-signal workers it has a probability of $p_L(2-p_L)$ of obtaining a good worker for its supervisory job in the second period. Its expected profits, π^p_{LL}, given a wage w^{LL} then are

$$\pi^p_{LL} = p_L\theta_G + (1-p_L)\theta_B - (p_L^2 + \frac{g}{1-g}p_H^2)(\bar{\theta}^F_G - \bar{\theta}^F_B) - 2w^{LL} \tag{14}$$

Substituting in for $\tilde{w}^{LL}$, this can only be an equilibrium if the firm is unwilling to employ the workers at this wage, which occurs if and only if

$$(1-p_L)^2(1-\chi)\bar{\theta}^W_B - (p_L^2 + \frac{g}{1-g}p_H^2)(\bar{\theta}^F_G - \bar{\theta}^F_B) \leq 0 \tag{15}$$

which is condition 2 of the proposition.

Finally, a non-promotion firm could, alternatively, employ two high-signal workers at wage w_H. The firm's expected profits, are simply two times the expected productivity of a high signal worker less wages. Therefore it must be the case that $\tilde{w}_H$ is less than the marginal product of the high-signal worker, which is true if and only if

$$-((1+\frac{g}{1-g})p_H^2 - 2(p_H - p_L))(\bar{\theta}^F_G - \bar{\theta}^F_B) - (1-a)^2(1-\chi)\bar{\theta}^W_B \leq 0 \tag{16}$$

which is the third condition of the proposition.

Taken together these four conditions assure that the wages w^{HH} and w_L, and factory price r, as defined above, represent a market clearing factory price and a characterize a set of contracts that both give non-negative benefits to all firms and workers and cannot be mutually improved upon. Since the parameters do not change over time, it is a steady-state equilibrium. ■

A.6 Proof of Corollary 10

Proof. This corollary follows directly from Proposition 9. The first two conditions are sufficient to make all the left-hand sides of Proposition 9's inequalities to be strictly increasing in θ_G. The right-hand side of the inequalities are unaffected by θ_G. Holding the right hand side fixed, there must by some θ_G large enough to make the left hand side's larger. The third condition just repeats condition 4 of Proposition 9, and guarantees that the equilibrium will be talent scarce. ■

A.7 Proof of Proposition 12

Proof. As shown in Proposition 9, the only observed wage at promotion firms for young entry-level workers, is

$$(p_H\theta_G + (1-p_H)\theta_B) + \frac{1}{2}(p_H(2-p_H) - (\frac{g}{1-g}p_H^2 + 2p_L)(\bar{\theta}^F_G - \bar{\theta}^F_B)) \tag{17}$$

Non-promotion firms pay entry-level workers their marginal product, and employ only low-signal workers, so

$$w_L = p_L\theta_G + (1-p_L)\theta_B \tag{18}$$

Taking the difference proves the proposition. ■

A.8 Proof of Proposition 13

Proof. The market for old entry-level workers is perfectly competitive, so all workers are paid their marginal product. The equilibrium is talent scarce, so no good workers are employed at entry-level jobs. Therefore, all old entry-level workers received the same wages, regardless of employer. ■

A.9 Proof of Proposition 14

Proof. Proposition 12 shows that entry-level workers are paid more when they are young at promotion than non-promotion firms. Proposition 13 shows that workers are paid identical wages when they are old at entry-level jobs, regardless of their firm's promotion policy. Subtracting the two wages proves the proposition. ■

A.10 Proof of Proposition 15

Proof. As derived in the proof of Proposition 9, the probability that a non-promotion firm is able to hire a good worker from the market is $\frac{g}{1-g}p_H^2 + 2p_L$. Condition 11 is simply that the probability that at least one of two randomly selected high-signal workers is good is higher than this. Since wages are determined by Nash bargaining, wages are identical at supervisory jobs conditional on worker type. ■

B Equilibrium Properties

B.1 Social Welfare Properties

Proposition 16 *Suppose an allocation of firms q and workers s is talent scarce. Then for any social welfare function S, there exists some set of feasible transfers t such that this set of worker and firm allocations solves the social planner's welfare optimization problem*

Proof. Suppose that there was a talent scarce equilibrium with allocation of workers s and firms q, but that there is a social welfare function S for which no set of transfers t will maximize given s and q. Then, there must be some other allocation of workers s' and firms q' that would allow for a feasible set of transfers t' that would provide a set of agents of positive measure with higher utility. Since all utilities are linear, this could only be the case if s' and q' increased the total output in the economy, thus allowing the size of t' to be larger than t. This would require that s' and q' increase the measure of the set of good workers assigned to supervisory jobs relative to s and q. However, by definition of talent scarcity, s and q assign the maximum measure to the set of good workers to supervisory jobs. Therefore s' and q' must not be feasible. ■

Proposition 17 *Suppose an allocation of firms q and workers s is not talent scarce and is not trivial. Then there does not exist any social welfare function S, such that a set of feasible transfers t would allows this set of worker and firm allocations to solve the social planner's welfare optimization problem*

Proof. Suppose there was some equilibrium with allocation of workers s and firms q that is not talent scarce, and some social welfare function, S, that is maximized by a feasible set of transfers t given s and q. Since the equilibrium is not talent scarce and not trivial, there is an alternative allocation of s' and q', namely setting the measure of promotion firms to 0 in periods that are not talent scarce, that will increase the measure of the set of good workers assigned to supervisory jobs, and thus the output of the economy. Since all utilities are linear, there must be a vector of transfers t' that under s' and q' that would increase the utility of all agents relative to s, q, and t. Therefore S is not a social welfare function. ■

These two propositions get at the general welfare properties of equilibria in this economy. While promotion-promise contracts can create an inefficiency in this economy in the sense that total production will be lower than in an economy where all firms do not commit, this is only the case when the equilibrium is not talent scarce. In a talent scarce equilibrium, the ability of firms to promote transfers utility from those who work at non-promotion firms to those who work at promotion firms, due to the latter's increased probability of being employed at a supervisory position. However, it still achieves the maximum output possible, since all old good workers are assigned to the more productive supervisory job. Therefore, a social planner with the ability to enforce lump sum transfers can do no better than adjusting these transfers. In a equilibrium that is not talent scarce, on the other hand, some old good workers are employed at entry-level jobs, while some old bad workers are employed at supervisory jobs. Output would be increased by switching the jobs of these workers, as would happen when no firms promote. A social planner could then use transfers to increase all workers utilities. I do not allow extra-wage contracts in my model. If I did, this would likely rule out equilibria that are not talent scarce via the first welfare theorem.

B.2 Other Pooling Equilibria

In this section of the appendix, I explore an alternative equilibrium in which each promotion firm hires one high-signal worker and one low-signal worker for the entry-level jobs. In this equilibria, promotion firms benefit by hiring a cheaper low-signal worker at a wage below his marginal product. The high-signal worker is then able to extract all of firm's future supervisory output, as opposed to in the two high-signal worker case, where that output was split in two. Non-promotion firms retain the strategy of hiring only low-signal workers. The proposition shows that, as long as this strategy yields a better supervisory worker on average than the non-promotion strategy, the testable predictions are the same as those for the equilibrium in Proposition 9.

Proposition 18 *Suppose that, in a talent scarce equilibrium, the set of firms who promotes is of measure* $2g$*, and each promotion firm hires one high-signal entry-level worker. Then, so long* $p_H - p_L - p_H p_L > \frac{2gp_Hp_L}{(1-2g)}$,

1. *Young entry-level workers at promotion firms are paid more on average than young entry-level workers at non-promotion firms*
2. *There will be no differences in average wages for old entry-level workers between promotion and non-promotion firms*
3. *The observed return to tenure of entry-level workers is higher at non-promotion firms*
4. *Supervisory job workers are paid more on average at promotion firms than at non-promotion firms*

Proof. First I will derive the wages that are paid at the promotion firm and the cost of factories. The excess demand for good supervisory workers at non-promotion firms is given by

$$1 - 2g > 2gp_Hp_L + 2(1-2g)p_L \tag{19}$$

where the first term on the right hand side is the measure of the set of good workers who are not promoted internally at promotion firms, and the second term is the measure of the set of good workers who were previously employed at non-promotion firms. This can be rearranged to show the probability of getting a good worker for a supervisory job at a non-promotion firm in this environment

$$\frac{2gp_Hp_L}{1-2g} + 2p_L \tag{20}$$

Since there are no profits in equilibrium (Lemma 2), and non-promotion workers are paid their marginal product (Lemma 5), the cost of factories must be the expected revenue at non-promotion firms from hiring supervisors from the market. Therefore

$$r = \bar{\theta}^F_B + (\frac{2gp_Hp_L}{1-2g} + 2p_L)(\bar{\theta}^F_G - \bar{\theta}^F_B) \tag{21}$$

Since low-signal workers are employed at both types of firms, they must be indifferent between the two sectors. Since workers are equally productive at entry-level jobs in all firms, and since talent is scarce so all that good workers are employed at supervisory jobs, the difference in the wages must be the difference in the probability of obtaining a supervisory position when the worker is revealed to be bad. Therefore, the wage for low-signal workers at entry-level promotion jobs, w^{LH}, is

$$w^{LH} = p_L\theta_G + (1-p_L)\theta_B - \frac{1}{2}(1-p_H)(1-p_L)(1-\chi)\bar{\theta}^W_B \tag{22}$$

where χ is the probability for a bad worker of obtaining a supervisory job at a non-promotion firm.

The expected profits of the promotion firm after paying factory costs, given they pay their high-signal workers w_{HL}, are

$$\pi^p_{HL} = (p_H+p_L)\theta_G+(2-p_H-p_L)\theta_B+(p_H-p_L-p_Hp_L-\frac{2gp_Hp_L}{(1-2g)})(\bar{\theta}^F_G-\bar{\theta}^F_B)-w^{HL}-w^{LH} \tag{23}$$

Substituting for w^{HL}, and setting profits equal to zero, the equilibrium high-signal entry-level wage is

$$w^{HL} = p_H\theta_G+p_L\theta_B+(p_H-p_L-p_Hp_L-\frac{2gp_Hp_L}{1-2g})(\bar{\theta}^F_G-\bar{\theta}^F_B)+\frac{1}{2}(1-p_H)(1-p_L)(1-\chi)\bar{\theta}^W_B \tag{24}$$

Each promotion firm hires one worker at w^{HL} and one worker at w^{LH}, so the average wage of a young entry-level worker at a promotion firm is

$$\bar{w}^p = \frac{1}{2}((p_H+p_L)\theta_G + (2-p_H-p_L)\theta_B) + \frac{1}{2}(p_H-p_L-p_Hp_L-\frac{2gp_Hp_L}{1-2g})(\bar{\theta}^F_G-\bar{\theta}^F_B) \tag{25}$$

The first term is unambiguously larger than the average wage for young workers at non-promotion firms, which is the expected marginal product of low-signal workers. The second term is greater than zero so long as $p_H - p_L - p_Hp_L > \frac{2gp_Hp_L}{(1-2g)}$, in which case the average wage for young entry-level promotion firms is larger than for the same jobs at non-promotion firms. Since talent is scarce this proves statements 1-3 (see proofs of Propositions 12-14).

The condition is that the probability of getting a good worker at a promotion firm is higher in equilibrium than for a non-promotion firm drawing a worker from the market. Since wages at supervisory jobs are determined by Nash bargaining and the firm type does not affect the threat points, this proves statement 4. ■

Table 1: Descriptive Statistics - Workers

	Entry-Level		Supervisorial	
	(1) Promotion	(2) Non-Promotion	(3) Promotion	(4) Non-Promotion
Tenure				
< 1 year	0.26	0.27	0.11	0.13
	(0.44)	(0.45)	(0.31)	(0.34)
1-2 years	0.16	0.17	0.12	0.13
	(0.37)	(0.38)	(0.33)	(0.33)
2-5 years	0.29	0.28	0.33	0.28
	(0.45)	(0.45)	(0.47)	(0.45)
5-10 years	0.17	0.14	0.21	0.21
	(0.38)	(0.35)	(0.41)	(0.41)
> 10 years	0.12	0.13	0.23	0.25
	(0.32)	(0.34)	(0.42)	(0.43)
Education				
None	0.18	0.23	0.08	0.14
GSCE	0.36	0.40	0.34	0.35
GCE	0.17	0.13	0.16	0.16
First Degree (BSc, BA, etc.)	0.24	0.18	0.32	0.28
Higher Degree (MSc, PhD, etc.)	0.06	0.05	0.11	0.08
Wage (Imputed)	8.76	7.75	12.76	11.27
	(3.99)	(3.54)	(5.28)	(5.15)
Occupation (1-digit SOC2000)				
Professionals	0.13	0.09	0.10	0.11
Technical Operations	0.19	0.16	0.12	0.13
Administrative	0.24	0.22	0.12	0.12
Skilled Trades	0.07	0.11	0.04	0.08
Personal Service	0.03	0.07	0.02	0.04
Sales	0.11	0.09	0.05	0.05
Process and Machine Operatives	0.13	0.12	0.05	0.03
Elementary Occupation	0.10	0.13	0.05	0.05
Managers			0.44	0.39
Observations	1798	3591	1334	2625

Source: WERS 2004 SEQ. Means and standard deviations are reported using weights provided with data.

Table 2: Descriptive Statistics - Employers

	(1) All	(2) Promotion	(3) Non-Promotion
Vacancy Policy			
Internal Only	0.00	0.01	0.00
	(0.06)	(0.12)	(0.00)
Internal Preferred	0.28	0.99	0.00
	(0.45)	(0.12)	(0.00)
No Preference	0.66	0.00	0.93
	(0.47)	(0.00)	(0.26)
External Preferred	0.01	0.00	0.01
	(0.09)	(0.00)	(0.10)
External Only	0.04	0.00	0.06
	(0.20)	(0.00)	(0.24)
Employees at Establishment	477.16	595.55	429.83
	(1157.40)	(1432.05)	(1024.27)
Employees in UK	8902.53	10465.74	8257.95
	(17430.74)	(18358.91)	(17000.89)
Establishment Age	41.21	41.28	41.19
	(60.69)	(55.48)	(62.68)
Percent Union Members	0.25	0.25	0.25
	(0.31)	(0.31)	(0.31)
Offers Incentive Pay	0.48	0.61	0.42
	(0.50)	(0.49)	(0.49)
Largest Occupation (1-digit SOC2000)			
Professionals	0.13	0.10	0.14
Technical Operations	0.14	0.09	0.15
Administrative	0.14	0.19	0.12
Skilled Trades	0.07	0.08	0.07
Personal Service	0.09	0.04	0.11
Sales	0.15	0.19	0.14
Process and Machine Operatives	0.14	0.17	0.13
Elementary Occupation	0.14	0.14	0.14
Observations	1533	422	1111

Source: WERS 2004 MQ. Means and standard deviations are reported using employment weights provided with data

Table 3: Non-Managerial, Non-Supervisorial Workers - Promotion Premium and Return to Tenure

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage
Promotion	0.136***	0.060***	0.102***	0.052***	0.041**
	(0.032)	(0.021)	(0.024)	(0.019)	(0.018)
Tenure	0.021***	0.020***	0.018***	0.013***	0.012***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Promotion * Tenure	-0.010	-0.008*	-0.012**	-0.009**	-0.008**
	(0.006)	(0.004)	(0.005)	(0.004)	(0.004)
Log Establishment Size					0.021***
					(0.004)
Worker Characteristics	No	No	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes	Yes
Industry FE	No	Yes	No	Yes	Yes
Geographic FE	No	No	No	No	Yes
Observations	5389	5389	5389	5389	5389
F-statistic	9.911***	4.044**	9.020***	3.931**	2.936*

Robust standard errors are clustered at the establishment level. F-statistic tests joint hypothesis of promotion dummy and tenure interaction. Worker characteristics include 6 education dummies, 16 ethnicity dummies, 8 age dummies, and a gender dummy. Occupation fixed effects use the UK Standard Occupation Classification 2000 (SOC2000) 3-digit system. Industry fixed effects use the UK Standard Industry Classification 2003 (SIC2003) 3-digit system. Geographic fixed effects are indicators for Government Office Region. Indicators are included for all top-coded variables, including an interaction between the tenure top-code and the promotion dummy

*p<.1, ** p<.05, *** p<.01

Table 4: Log Wage Regressions for Non-Managerial, Non-Supervisorial Workers with 10 or More Years Seniority

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage
Promotion	0.138**	0.030	0.023	0.027
	(0.058)	(0.037)	(0.040)	(0.039)
Log Establishment Size				0.025**
				(0.011)
Worker Characteristics	No	No	Yes	Yes
Occupation FE	No	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes
Geographic FE	No	No	No	Yes
Observations	691	691	691	691

Robust standard errors are clustered at the establishment level. Worker characteristics include 6 education dummies, 16 ethnicity dummies, 8 age dummies and a gender dummy. Occupation fixed effects use the UK Standard Occupation Classification 2000 (SOC2000) 3-digit system. Industry fixed effects use the UK Standard Industry Classification 2003 (SIC2003) system. Geographic fixed effects are indicators for Government Office Region. Indicators are included for all top-coded variables.

* p<.1, ** p<.05, *** p<.01

Table 5: Log Wage Regressions for Managerial and Supervisorial Workers

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage	(5) Log Wage
Promotion	0.074***	0.020	0.009	0.032	0.001
	(0.025)	(0.018)	(0.016)	(0.031)	(0.028)
Tenure				0.008**	0.006**
				(0.003)	(0.003)
Promotion * Tenure				-0.002	0.004
				(0.006)	(0.005)
Log Establishment Size			0.029***		0.028***
			(0.005)		(0.005)
Worker Characteristics	No	No	Yes	No	Yes
Occupation FE	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Geographic FE	No	No	Yes	No	Yes
Observations	3959	3959	3959	3959	3959

Robust standard errors are clustered at the establishment level. Worker characteristics include 6 education dummies, 16 ethnicity dummies, 8 age dummies, and a gender dummy. Occupation fixed effects use the UK Standard Occupation Classification 2000 (SOC2000) 3-digit system. Industry fixed effects use the UK Standard Industry Classification 2003 (SIC2003) 3-digit system. Geographic fixed effects are indicators for Government Office Region. Indicators are included for all top-coded variables.

* p<.1, ** p<.05, *** p<.01

Table 6: Ordered Logit Regressions on Highest Educational Attainment

	(1) All Workers	(2) All Workers	(3) Entry-Level	(4) Supervisory
Promotion	0.338***	0.120*	0.151*	0.070
	(0.095)	(0.063)	(0.082)	(0.092)
Log Establishment Size		0.120***	0.081***	0.179***
		(0.023)	(0.028)	(0.034)
Worker Characteristics	No	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes
Industry FE	No	Yes	Yes	Yes
Geographic FE	No	Yes	Yes	Yes
Observations	9348	9348	5389	3959

Robust standard errors are clustered at the establishment level. Worker characteristics include 16 ethnicity dummies, 8 age dummies, and a gender dummy. Occupation fixed effects use the UK Standard Occupation Classification 2000 (SOC2000) 3-digit system. Industry fixed effects use the UK Standard Industry Classification 2003 (SIC2003) 3-digit system. Geographic fixed effects are indicators for Government Office Region. Indicators are included for all top-coded variables.

* p<.1, ** p<.05, *** p<.01

Table 7: Non-Managerial Non-Supervisoral Workers - Establishment Productivity Measures

	(1) Log Wage	(2) Log Wage	(3) Log Wage
Promotion	0.027	0.028	0.060*
	(0.035)	(0.029)	(0.032)
Tenure	0.014***	0.013***	0.015***
	(0.005)	(0.004)	(0.005)
Promotion * Tenure	-0.013**	-0.013**	-0.015**
	(0.006)	(0.005)	(0.006)
Log Establishment Size	0.040***	0.038***	0.021***
	(0.008)	(0.007)	(0.007)
Log Profit / Worker	0.005		
	(0.005)		
Log Value Added / Worker		0.013*	
		(0.008)	
Log Capital / Worker			0.020**
			(0.010)
Worker Characteristics	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes
Observations	1573	2226	1791
F-statistic	2.336*	3.264**	3.071**

Robust standard errors are clustered at the establishment level. *F*-statistic tests joint hypothesis of promotion dummy and tenure interaction. Worker characteristics include 6 education dummies, 16 ethnicity dummies, 8 age dummies, and a gender dummy. Occupation fixed effects use the UK Standard Occupation Classification 2000 (SOC2000) 3-digit system. Industry fixed effects use the UK Standard Industry Classification 2003 (SIC2003) 3-digit system. Indicators are included for all top-coded variables.including an interaction between the tenure top-code and the promotion dummy.

* p<.1, ** p<.05, *** p<.01