BONUSES AND PROMOTION TOURNAMENTS: THEORY AND EVIDENCE*

Forthcoming in the Economic Journal

Emre Ekinci, Antti Kauhanen and Michael Waldman

Abstract

Standard models of promotion tournaments do not distinguish between wages and bonuses and thus cannot explain variation in the use of bonuses. We combine classic and market-based tournament theories to develop a model in which wages and bonuses serve distinctly different roles. We use this model to derive testable predictions which we test employing both a single firm dataset and a dataset encompassing a large segment of the Finnish economy. Our empirical analysis supports the testable predictions and shows that our theoretical approach better matches the data than alternative theories of bonus determination based on arguments already in the literature.

A seminal contribution in the personnel economics literature is the idea of promotion tournaments first put forth in Lazear and Rosen (1981). In this theory large wage increases are attached to promotions in order to achieve efficient effort levels for lower level workers. One drawback of the Lazear and Rosen approach, however, is that it makes no distinction between

---

* Corresponding author: Emre Ekinci, Universidad Carlos III de Madrid, Calle Madrid, 126, 28903 Getafe (Madrid), SPAIN. Email: emre.ekinci@uc3m.es.

We thank Fran Blau, Kevin Hallock, George Jakubson, Victoria Prowse, Kjell Salvanes (the editor), Jan Zabojnik, two anonymous referees, and seminar and conference participants at Cornell University, MIT, Universidad Carlos III de Madrid, Queen's University, IAE-CSIC, and the 2016 SOLE meetings for helpful comments. We also thank Michael Gibbs for providing one of the datasets used in this study. Financial support from Spain's Ministry of Economy and Competitiveness (ECO2016-75961-R) and the Academy of Finland (Grant no. 258771) are gratefully acknowledged.

wage increases and bonuses and thus cannot be used to explain how bonus payments vary across workers and job levels inside a firm. In this paper we combine the classic approach to promotion tournaments pioneered by Lazear and Rosen with the market-based approach to promotion tournaments first analysed in Gibbs (1995) and Zabojnik and Bernhardt (2001) to develop a theory concerning the role of bonuses in promotion tournaments.

In the classic approach to promotion tournaments pioneered by Lazear and Rosen (1981) each firm commits to a high wage associated with promotion, a low wage for workers not promoted, and also commits to promote the worker who produces more output. The result is that the high promotion wage serves as an incentive for effort and by optimally choosing the wage spread the firm induces low level workers to choose efficient effort levels. In contrast, in the market-based approach to promotion tournaments first explored in Gibbs (1995) and Zabojnik and Bernhardt (2001), firms do not commit to high promotion wages. Rather, building on Waldman (1984a), the promotion serves as a signal of high worker productivity. In turn, the signal results in high wage offers for promoted workers from prospective employers and the initial employer responds with a high promotion wage in order to stop promoted workers from being bid away. Like in classic tournament theory, the high wage serves to increase incentives for low level workers.[1]

In this paper we construct a model that combines features of the two approaches. Like in the market-based approach, firms cannot commit to the size of compensation increases associated with promotion, but rather the size of these increases is determined by the signal associated with promotion and the resulting higher compensation a promoted worker would receive by moving to an alternative employer. Like in the classic approach, on the other hand, we do allow firms to have some commitment ability in terms of compensation increases. In particular, at the beginning of each period each firm commits to a rate at which bonus size

[1] See Prendergast (1999) and Lazear and Oyer (2013) for surveys that discuss the classic tournament approach and Waldman (2013a,b) for surveys that discuss both the classic and market-based approaches.

increases with worker output. In our model, as in earlier market-based tournament models, incentives provided through promotion prizes are typically not first best. Thus, the role of the bonus is to augment promotion prizes so that aggregate incentives for effort are first best.

One empirically well documented finding concerning bonuses is that the size of bonus payments increases with job level (see, for example, Lambert *et al*. (1993), Baker *et al*. (1994a,b), and Smeets and Warzynski (2008)). However, the economic rationale behind this finding is not well understood. We show that our model captures this finding and, more generally, generates five testable predictions concerning how bonuses should vary across individuals and across jobs within a firm. These testable predictions are: i) controlling for age and job level tenure, bonus payments increase with job level; ii) controlling for job level and age, bonus payments increase with job level tenure; iii) controlling for job level, and job level tenure, bonus payments increase with worker age; iv) controlling for age, job level, and job level tenure, bonus payments increase with performance; and v) the bonus payment in the current period is negatively related to the expected prize associated with future promotion.

These five predictions all follow from the idea that in our model the firm always chooses the bonus rate that achieves efficient effort choices, where efficiency requires equality between the worker's marginal cost of effort and the marginal benefit associated with additional effort. This idea translates into our five testable predictions as follows. First, related to an argument in Rosen (1982), in our model the incremental productivity associated with additional effort increases with job level, so achieving efficient effort levels requires that the incremental compensation associated with additional effort also increases with job level. This is the primary driver of our first prediction that, holding job level tenure and age fixed, the size of bonuses increases with job level. Second, due to an assumption of task specific human capital, the increasing productivity associated with additional effort increases with job level tenure. So the bonus also increases with job level tenure. Third, because of the accumulation of human capital as workers age, the increasing productivity associated with added effort increases with age and thus so do bonus payments. In addition, the reduction in promotion incentives as workers age also contributes to this result. Fourth, the prediction that bonus size rises with performance follows immediately as long as the bonus rate is positive which is the case in our model as long as promotion incentives are not too high. Fifth, if the expected prize for promotion rises, then bonus size must fall for overall incentives to remain at the efficient level.

In the second part of the paper we test these predictions. Our first set of tests uses personnel data from a medium-sized firm in the financial services industry. This dataset was first employed in the classic study of Baker *et al*. (1994a,b) that provides a detailed examination of wage and promotion dynamics at the firm.[2] The original dataset was a twenty-year unbalanced panel consisting of all managerial employees at the firm. We only employ the last seven years of the dataset in which bonus information is available. This seven-year panel is well suited for our purposes because in addition to having information on salary, bonus, and performance ratings, it also includes detailed information on the firm's job ladder that Baker, Gibbs, and Holmstrom constructed using the raw data on job titles and typical movements across job titles.

Our empirical analysis of this dataset provides support for most of the model's predictions. Specifically, regression results show that the size of bonus payments increases with job level even when we control for age and job level tenure, the size of the bonus increases with job level tenure controlling for job level and age, and the size of the bonus payment also increases with performance after controlling for age, job level, and job level tenure. Consistent with the fifth prediction, we also find that an increase in the expected prize associated with promotion in the following period is negatively correlated with the size of the current bonus. In other words, as captured in our theoretical model, we find a trade-off between explicit incentives from bonuses and implicit incentives that arise from the tournament aspect of promotions. We do not find clear support, however, for the prediction that bonus size should increase with age

---

[2] Other studies that employ this dataset include Gibbs (1995), DeVaro and Waldman (2012), and Kahn and Lange (2014).

controlling for job level and job level tenure. Below we discuss one possible reason for why this prediction has limited support in this dataset.

Our second set of tests employs data from a large linked employer-employee dataset from Finland that is representative of white-collar employees in manufacturing. We use data from the time period 2003 to 2012 because of a change in the classification of job titles in the previous year.[3] This dataset includes most of the same information as found in the Baker *et al*. (1994a,b) dataset except that it does not include any information concerning performance evaluations. Using this dataset we reproduce all of the earlier tests that do not include performance evaluations and we find similar results. Specifically, we find that bonus payments increase with job level controlling for age and job level tenure, increase with job level tenure controlling for job level and age, and that bonus payments decrease with the expected prize associated with promotion. Further, in this dataset we find some support for bonus size increasing with age at lower age levels after controlling for job level and job level tenure.

Later in the paper we discuss in detail the mixed support we find for the theoretical prediction concerning age. But basically we feel our findings concerning the age prediction are more supportive of the theoretical prediction than they might at first seem. First, the theoretical prediction is really that bonus size should increase with age up to the age at which human capital accumulation peaks. Findings concerning the Finnish economy suggest that for this economy this peak occurs when workers are in their early to mid 50s. This is basically consistent with some of our regressions concerning the Finnish dataset where we find that bonus size rises with age up to workers being between their mid 40s and mid 50s. Second, given the likely importance of stock holdings and stock options in the Baker, Gibbs, and Holmstrom dataset (but not in the Finnish dataset) and that we do not have measures of stock holdings and stock options in that dataset, it is not surprising that the age prediction concerning bonus size does not find consistent support in our tests employing the Baker, Gibbs, and Holmstrom dataset.

---

[3] This dataset has been investigated in a number of previous studies including Kauhanen and Napari (2012) which shows that empirical patterns in this dataset are similar to those found in Baker *et al*. (1994a,b).

Note that our empirical analysis suggests that the theoretical model developed in this paper better matches the data concerning how bonuses vary across job levels and individuals within firms than competing theories based on arguments already in the literature. One competing theory that builds on Rosen (1982) and Lemieux *et al*. (2009) is that bonuses are solely driven by how the return to effort varies across job levels. This theory does not explain the evidence we find in our empirical analysis of the Finnish dataset concerning bonuses rising with age at younger ages after controlling for job level and job level tenure. Another competing theory based on the analysis in Gibbons and Murphy (1992) is that bonuses are solely driven by decreasing career concern incentives as workers age, but this theory does not explain why bonuses rise with job level after controlling for worker age. And neither of these alternative theories explains our findings of a positive correlation between bonus size and job level tenure and a negative correlation between bonus size and promotion based rewards.

The outline for the paper is as follows. Section 1 discusses related literature. Section 2 presents our model and some preliminary results. Section 3 presents an analysis of the full equilibrium of the model and discusses testable predictions. Section 4 describes the data we use in our empirical analysis. Section 5 begins with a preliminary empirical analysis of the firm's bonus policy and then presents our investigation of the model's testable predictions. Section 6 presents concluding remarks.

1. Related Literature

As briefly discussed in the Introduction, our paper falls into the extensive literature on promotion tournaments which started with the seminal contribution of Lazear and Rosen (1981).[4] All of the early literature on the subject assumes that the incentive effects of promotions stem from an ability of firms to commit to future compensation levels. That is, firms commit to high levels of compensation for promoted workers and lower levels for workers not promoted and

[4] Other early papers in this literature include Green and Stokey (1983), Nalebuff and Stiglitz (1983), Malcomson (1984), and O'Keefe *et al*. (1984).

then workers compete for promotion prizes typically through the choice of effort levels. This literature considers a variety of analyses which include comparing promotion tournaments with other ways of compensating workers, deriving the properties of equilibrium promotion tournaments under various assumptions concerning worker heterogeneity, the nature of the production environment, etc., and considering multi-stage promotion tournaments.

A more recent literature that has come to be called the market-based approach assumes that commitment is not possible, but rather promotion prizes arise due to the signalling role of promotions first explored in Waldman (1984a). The basic argument, first put forth in Gibbs (1995) and Zabojnik and Bernhardt (2001), is that firms pay promoted workers high promotion wages in order to stop promoted workers from being bid away when the positive signal associated with promotion results in prospective employers increasing their wage offers.[5] In turn, like in the Lazear and Rosen approach, workers respond to the high promotion wages by increasing effort or investing more in the development of human capital.[6]

Our model combines elements of each of these two approaches. We allow firms some commitment ability in terms of compensation – in each period firms commit to a minimum output required to receive a bonus and a bonus rate for that period. However, like in the market-based approach, firms cannot commit to compensation levels for future periods and promotion prizes arise due to the signalling role of promotion rather than commitment. Further, different

---

[5] See Ghosh and Waldman (2010), Zabojnik (2012), and Gürtler and Gürtler (2015) for more recent papers that take this approach.

[6] A small literature has developed that focuses on whether real world promotion tournaments are better described by the classic approach put forth by Lazear and Rosen (1981) or the more recently developed market-based approach. Waldman (2013b) surveys the empirical literature and argues that the evidence is mixed concerning whether it is more consistent with the classic approach or the market-based approach. He concludes by arguing that a hybrid approach that combines the two approaches which is similar to the approach we pursue might be more consistent with the evidence than either of the two approaches taken in their pure forms. Wang (2013) focuses on predictions concerning how promotions affect turnover and, based on an empirical analysis of the Baker *et al.* (1994a,b) dataset, argues that the market-based approach does a better job of explaining the evidence at low levels of the job ladder and the classic approach does a better job at high levels. In contrast, DeVaro and Kauhanen (2016) provide empirical tests based on how worker and firm behavior changes when the stochastic component of worker performance becomes more important. They find that the classic approach does best at matching results based on this set of tests.

than both approaches, our focus is on the role and size of bonuses in promotion tournament settings which has drawn little prior attention in this literature.[7]

Although not promotion tournament models, the paper also builds on Rosen (1982), Gibbons and Murphy (1992), and Lemieux *et al*. (2009). Gibbons and Murphy extend the career concerns argument of Holmstrom (1999) (see also Fama (1980)) to consider how performance pay should vary over a worker's career. The basic argument is that as a worker gets older incentives provided through the symmetric learning mechanism identified by Holmstrom should decrease. This occurs both because of fewer periods remaining in the worker's career to reap any return from improved beliefs concerning the worker's ability and because there is less remaining uncertainty concerning the worker's ability. As a result, to keep incentives high as workers age, firms must increase pay for performance. After developing the theory, Gibbons and Murphy show supporting evidence using data on CEO compensation.

Our argument is related in that we also focus on a trade-off between different avenues through which incentives are provided, although the specific avenues considered are different across the two papers. Gibbons and Murphy (1992) focus on career concern incentives and performance pay (bonuses are a type of performance pay), while our focus is promotion incentives and bonuses. Further, in addition to this difference in focus, there are a number of other important differences. First, they assume firms learn about worker ability after they enter the labour market in a symmetric fashion which means that all firms have the same information about each worker's ability at any point in time. In contrast, we assume this learning is asymmetric which means that a worker's current employer is better informed. Second, they

---

7 There are a few earlier papers that consider the possibility of a trade-off between promotion-based incentives and bonuses. Gibbs (1995) and Boschmans (2008) both provide theoretical analyses that capture the basic trade-off, although neither provides results similar to any of our other testable predictions. Gibbs also provides empirical testing but finds no evidence of a trade-off between promotion-based incentives and bonuses, while Boschmans provides no empirical testing. Ortin-Angel and Salas-Fumas (1998) provide an empirical investigation of bonuses in a sample of top and middle managers at Spanish firms and finds a number of results some of which they interpret as evidence of a trade-off between promotion-based incentives and bonuses. But they do not test for the trade-off directly and their findings, such as that bonuses rise with job level, have alternative explanations. Also, Krakel and Schottner (2008) consider a theoretical model characterized by promotion tournaments and bonuses, but their model does not capture the trade-off between promotion-based incentives and bonuses.

assume a single job at which a worker remains throughout his or her career, while we assume a job ladder that workers climb as they gain labour market experience and show evidence of superior ability through successful performance. Third, in addition to a prediction similar to the Gibbons and Murphy prediction concerning how pay for performance varies with worker age, we derive testable predictions concerning bonus size, job level, job level tenure, and promotion incentives not analogous to anything found in that earlier paper.

Lemieux *et al*. (2009) show how the prevalence of pay for performance affects wage inequality. Although not their main focus, their theoretical model suggests a potential explanation for why bonuses increase with job level. That is, in their model the size of performance pay at a job depends positively on the returns to effort at that job. Combining this idea with the one found in Rosen (1982) that returns to effort are higher at higher job levels yields the prediction that the performance pay component of compensation contracts should be larger at higher levels of a firm's job ladder.[8] We incorporate this idea into our model and it is an important driving force behind our testable predictions.

Note that Gibbons and Murphy (1992) and the combination of Rosen (1982) and Lemieux *et al*. (2009) are two alternative explanations for why bonuses rise with job level. First, in the Gibbons and Murphy (1992) argument bonuses are predicted to rise with job level because, on average, workers on higher job levels are older. Thus, in that argument, holding worker age fixed, bonuses are independent of job level, but bonuses should rise with age holding job level fixed. Second, in the combined argument of Rosen (1982) and Lemieux *et al*. (2009) bonuses rise with job level because the return to worker effort rises with job level. In this argument bonuses should rise with job level even holding worker age fixed, but bonuses are predicted to be independent of worker age holding job level fixed. In our argument, in contrast, bonuses rise with job level holding worker age fixed and rise with age holding job level fixed.

[8] To be precise, in Rosen (1982) returns to worker ability are higher at higher job levels (see also Waldman (1984b)). But if productivity is a function of ability plus effort (which is what we assume in our theoretical analysis) or ability times effort, then returns to worker ability rising with job level will also translate into returns to worker effort rising with job level.

Further, our model predicts a positive correlation between job level tenure and bonus size and also a trade-off between bonus incentives and promotion incentives. Neither of these additional predictions is generated by these alternative theories.[9]

Our assumption that the labour market is characterized by asymmetric learning and promotion signalling is supported by a number of empirical studies. Gibbons and Katz (1991) were the first to empirically test for asymmetric learning in labour markets. Their focus was Greenwald's (1986) adverse selection argument concerning labour market turnover and its implications for differences between laid off workers and those fired in a plant closing. Using the Current Population Survey (CPS), they find support for the adverse selection argument. Further, a number of more recent papers including Schönberg (2007), Pinkston (2009), and Kahn (2013) take alternative approaches to test for asymmetric learning and, in general, these more recent papers also find evidence consistent with asymmetric learning being important.[10]

There are also papers that directly consider the promotion-as-signal hypothesis. Most of these papers focus on tests derived from the basic idea first put forth in Bernhardt (1995) that the signal associated with promotion should be smaller for workers in higher education groups, so workers in these groups should be favoured in the promotion process. The papers that take this approach such as Belzil and Bognanno (2010), DeVaro and Waldman (2012), Cassidy *et al.* (2016), and Bognanno and Melero (2016) mostly find evidence that supports the hypothesis.[11]

We also assume that the production process is characterized by task specific human capital which is the focus of a number of recent theoretical and empirical papers. The basic definition of task specific human capital and discussions of various potential applications can be found in Gibbons and Waldman (2004,2006). Empirical studies that support the task specific

---

[9] Benabou and Tirole (2016) recently put forth a new theory of bonus pay based on screening and multi-tasking but it does not explain the empirical findings we are focused on such as the evidence that bonuses rise with job level.

[10] Of these more recent papers, Schönberg's is the only one that finds weak evidence for asymmetric learning. But as is argued in Waldman (2013a), it is unclear that the test for which she finds weak evidence for asymmetric learning is in fact a valid test of the asymmetric learning argument.

[11] Gibbs (2003) employs alternative tests of the promotion signaling hypothesis that do not depend on how promotion signals vary with education and he also finds evidence that supports the hypothesis. Also, see Dato *et al.* (2016) for an experimental labor market analysis that supports the promotion signaling argument.

human capital argument include Gathmann and Schönberg (2010), Schulz *et al.* (2013), Cook and Mansfield (2016), DeAngelo and Owens (2017), and Jin and Waldman (2017).

The paper is also related to a well known puzzle identified by Baker *et al.* (1988). That paper asked, why is it that incentives are provided through promotions rather than solely through bonuses and other non-promotion based compensation increases? That is, if promotions are used to both assign workers to jobs and provide incentives, then inefficiencies will result because the two roles will sometimes be in conflict. So, according to Baker, Jensen, and Murphy, it would be more efficient to use promotions solely for assignment and use non-promotion based compensation changes to provide incentives. The market-based approach to promotion tournaments provides an answer to this puzzle. Specifically, the signalling role of promotions causes promotions to be associated with large wage increases, so firms are in a sense forced to employ promotions as an incentive device. In this paper we extend this argument to consider how the provision of incentives is divided between bonuses and promotion-based incentives when promotions serve as signals.

2. Model and Preliminary Analysis

In this section we present our model of promotion and bonus incentives in a hierarchical model of production, and then present some preliminary results.[12] In the next section we present a full equilibrium analysis and also present and discuss testable implications. Note that the specific model we consider builds on an analysis in Ghosh and Waldman (2010).

[12] In the model we construct and analyze there is no requirement that only a single worker or some fixed number of workers is promoted. So one might argue that this is not a promotion tournament model. Following Waldman (2013b), we are defining promotion tournaments as settings in which promotions serve an incentive role because of promotion wage increases whether or not there is a requirement that only a single worker or a fixed number of workers is promoted. Note also that adding a limit on the number of workers a firm can promote would complicate the analysis without changing the qualitative nature of the results. See Waldman (2013b) for an example of a market-based tournament model in which each firm can promote only a single worker.

2.1 The Model

We consider a two-period model with free entry and identical firms that produce output using labour as the only input. Workers live two periods, where a worker is referred to as young in period 1 and old in period 2. Worker i's innate ability is denoted θ_i, where there are two groups of workers denoted groups 1 and 2. The value for θ_i for a worker in group k is a random draw from the probability distribution function $F_\theta^k(.)$ with support $[\theta_k^L,\theta_k^H]$, where $E(\theta^2)-\theta_2^L \geq E(\theta^1)-\theta_1^L>0$ and $E(\theta^k)$ denotes the unconditional expected value for θ for a worker in group k. None of the labour market participants including the worker knows a worker's true value for θ_i at the beginning of the game but a worker's group as well as the distributions $F_\theta^1(.)$ and $F_\theta^2(.)$ are common knowledge. The asymmetric information structure of the model determines how firms learn about worker ability. A worker's current employer and the worker privately observe the worker's output realization at the end of each period and use that information to revise beliefs about the worker's ability. The firm then uses this additional information in deciding whether or not to promote the worker while prospective employers use the promotion decision as a signal of ability.

Each firm has three job levels, denoted 1, 2, and 3. If worker i is assigned to job j, j=1,2, or 3, in period t, then the worker produces $y_{ijt}=s_{it}(c_j+d_j(\theta_i+e_{it}))-z_{ijt}$, where s_{it} is worker i's human capital in period t which is defined in detail below, e_{it}, $e_{it}\geq 0$, is worker i's effort in period t, and z_{ijt} represents a training cost which is also described in detail below. Note that, given no stochastic element in the production process, at the end of the first period in any pure strategy Nash equilibrium each firm learns with certainty the innate ability levels of its period 1 employees. Introducing a stochastic element would complicate the analysis, but would not change the basic nature of the results.

Starting with Rosen (1982), it is standard to assume that the incremental productivity associated with ability (and effort) increases with job level and that it is efficient to assign low ability workers to low levels of the job ladder and high ability workers to high levels of the job ladder. To make the model consistent with these standard conditions we assume $d_3>d_2>d_1>0$ and $0<c_3<c_2<c_1$. Also, additional related assumptions are imposed below.

As indicated, s_{it} is worker i's human capital in period t. For all young workers s_{it} equals one. If worker i does not switch firms at the beginning of period 2, then $s_{i2}=s_1$ if this is the worker's first period at the worker's current job level and $s_{i2}=s_2$ if it is the second period on the current level, where $s_2>s_1>1$. If at the beginning of period 2 worker i moves to a new firm, then $s_{i2}=h_1$ if this is the worker's first period at the worker's current job level and $s_{i2}=h_2$ if it is the second period on the current level, where $h_2>h_1>1$, $s_2>h_2$, and $s_1>h_1$. In other words, in our model there is general, firm specific, and task specific human capital. General human capital is measured by h_1-1, while firm specific human capital is captured by the differences s_1-h_1 and s_2-h_2. The task specific human capital some of which is also firm specific is captured by the differences s_2-s_1 and h_2-h_1. As discussed in Gibbons and Waldman (2004), task specific human capital can be thought of as the accumulation of human capital which is partially or fully lost when a worker switches tasks, which we capture by a switch in the job level. Note that a worker can accumulate general, firm specific, and task specific human capital either through learning-by-doing or on-the-job training. Since it is not our focus, the exact mechanism through which workers accumulate the various types of human capital that we assume is not modelled.

We also assume $s_1d_{k+1}>s_2d_k$ and $h_1d_{k+1}>h_2d_k$ for all k, k=1,2. Since our focus will be on parameterizations such that group k workers are on job k in period 1, these assumptions tell us that for old workers productivity rises faster with innate ability after a promotion. The role of these assumptions is to help ensure that firms have an incentive to place workers with higher innate ability at higher levels of the job ladder. Additionally, s_1d_{k+1}-$s_2d_k>h_1d_{k+1}$-h_2d_k for all k, k=1,2. This assumption states that the net increase in productivity associated with promotion rises faster with innate ability when the promotion does not include a move between firms. This assumption also helps ensure that firms promote higher innate ability workers.

Employing data from the Panel Study of Income Dynamics (PSID), Brown (1989) finds results consistent with the existence of positive and substantial training costs associated with starting work in a new position, where these costs depress both productivity and compensation during the training period. Also, in Brown's analysis, a new position is defined as either a new job assignment for a worker who does not switch employer or any job assignment immediately after a worker switches employers.

We incorporate this idea into our model through the training cost, z_{ijt}, which denotes the training cost worker i incurs if assigned to job j in period t. Specifically, there is a positive training cost the first period a worker is at a firm and the second period a worker is at a firm but it is the first period at the current job level. So $z_{ijt}=z$, $z>0$, if $t=1$, $z_{ijt}=z$ if $t=2$ and worker i was not on job level j in period 1, and $z_{ijt}=z$ if $t=2$ and worker i is at a new firm in period 2. Otherwise, $z_{ijt}=0$. Note that we further assume that the training cost is sufficiently large that the model exhibits a promotion signalling distortion, and also that $c_1>z$ which ensures that output is always positive. We provide a further discussion of the role of the training cost after Proposition 1 which describes the full equilibrium of the model. See also the proof of Proposition 1 in Appendix A.

In our specification in which a worker's ability and effort are added together in the production function the first best efficient effort level for a worker is independent of the worker's ability, holding the worker's job assignment and human capital fixed. This aspect of our specification is not essential and, in fact, basically all of our results would continue to hold if we assumed, for example, that output was a function of ability times effort rather than ability plus effort. In that case the two would be complementary meaning that first best effort, holding a worker's job assignment and human capital fixed, would be positively related to worker ability. We employ the specification in which the two are added together rather than multiplied because this simplifies the analysis and thus makes the logic of our arguments easier to follow. It is similarly the case that the assumed complementarity between ability and human capital is not essential.

In our specification effort and human capital are also complementary meaning that, holding a worker's job assignment fixed, first best effort is positively related to the worker's human capital. Many of our theoretical predictions depend on this aspect of our specification. For example, due to the accumulation of task specific human capital, holding everything else fixed, first best effort rises with a worker's job level tenure that, in turn, means the expected bonus rises with job level tenure. There are results in the empirical literature consistent with our assumption that human capital and effort are complementary.[13]

We assume that employers cannot offer long-term contracts, i.e., they cannot commit to future wages or promotion decisions in subsequent periods. Also, prospective employers do not observe salaries paid or bonuses. In order to capture the interaction between the size of bonus payments and promotion incentives, we also assume that output is contractible but not publicly observable.[14] To be specific, in each period t, t=1,2, firms offer workers compensation contracts consisting of a base salary, α, a minimum output level, y^M, and a bonus rate, β, where the bonus payment is the bonus rate multiplied by the difference between the worker's output and the minimum output level specified in the contract.

Firms and workers are assumed to be risk neutral and both have a discount factor equal to δ, where δ is sufficiently small that the bonus rate is always positive. Further, worker utility is given in equation (1).

$$U(w_1,w_2,e_1,e_2)=\sum_{t=1}^{2} \delta^{t-1}U_t(w_t-g(e_t))= \sum_{t=1}^{2} \delta^{t-1}(w_t-g(e_t)) \tag{1}$$

In this equation w_t is the worker's wage in period t (salary plus bonus), e_t is the period t effort level, and $g(e_t)$ is the disutility of effort. We further assume $g(0)=0$, $g'(0)=0$, $g'(e)>0$ and $g''(e)>0$ for all $e>0$. Let $e_j^*(s_{it})$, j=1,2, be the efficient effort choice for worker i in period t assigned to

[13] For example, Kuhn and Lozano (2008) study the determinants of which types of workers work long hours (more than 48 hours per week) and find a number of results indicating that working long hours is more common for workers with high levels of human capital. These results are consistent with our specification if we interpret high hours per week as an indication of a high effort level.

[14] In other words, the employer can credibly reveal a worker's output to the courts but that output is not observable to prospective employers. See Mukherjee (2008) and Koch and Peyrache (2009) for papers that employ this approach. An alternative approach taken by some authors is to assume that a supervisor, whose responsibility includes assessing worker performance, can bias assessments in order to misallocate monetary rewards. See, for example, Milgrom and Roberts (1988), Prendergast and Topel (1996), and Fairburn and Malcomson (2001) for related analyses.

job j. Given the specification for production and worker utility described above, $e_j^*(s_{it})$ satisfies $s_{it}d_j=g'(e_j^*(s_{it}))$ for j=1,2, and 3.[15]

An alternative theoretical approach would be not to assume that δ is sufficiently small that the bonus rate must be positive, but assume instead that the bonus rate cannot be negative, for example, for legal reasons. The only difference in that specification is that there would be allowable parameterizations for which the bonus rate for young workers on a particular level could equal zero and also some of the theoretical results would not hold for some of these workers. For example, one of our findings is that the effort choice is always at the efficient level. In this alternative specification, however, there could be young workers with zero bonus payments and for these workers effort could actually exceed the efficient level. The logic is that the worker earns a zero bonus because promotion incentives exceed efficient incentives, so efficient effort is only achieved with a negative bonus. If in that case the bonus is not allowed to be negative, the result is that effort exceeds the efficient level. Note that one way to interpret the analysis and results that follow is that the results all hold for this alternative specification for all workers for whom the bonus rate is strictly positive.

The sequence of moves in the game is as follows. At the beginning of period 1 firms make period 1 contract offers and workers choose firms based on utility maximization. Workers are then assigned to jobs and each worker chooses an effort level. Then output is produced, privately observed by the first period employer and the worker, and then the worker is paid. Based on the output realization, each worker's first period employer updates its beliefs regarding the worker's ability and then, at the beginning of period 2, assigns the worker to a job. Prospective employers, which observe neither effort nor output, use the job assignment as a signal of ability and make offers consisting of a compensation contract and job assignment. Then each worker's current employer is allowed to make a counter-offer concerning the compensation

[15] We further assume that g'' in the relevant range, i.e., the range associated with the various values for efficient effort, is large which ensures that the disutilities for effort associated with these efficient effort levels are not far apart. This helps ensure that in equilibrium firms always have an incentive to promote higher ability workers.

contract. Each worker then chooses a firm, chooses a second period effort level, second period output is realized, and the worker is paid.

Our specification of the timing of job assignments and compensation offers is a standard approach in this literature since Bernhardt (1995). One positive aspect of this approach is that assuming that the incumbent can make a counter-offer after prospective employers make their compensation offers is consistent with empirical evidence in Barron *et al*. (2006). It is the case, however, that in our model results are robust to assuming various alternatives concerning the timing of compensation offers. For example, results would be the same under the assumption that the incumbent firm and prospective employers make alternating offers as long as the incumbent firm moves last. Results would also be the same under the assumption that the incumbent and prospective employers make simultaneous compensation offers and the incumbent is not allowed to make a counter-offer. See footnote 19 for further discussion of this issue.

Our focus is on pure strategy Perfect Bayesian equilibria where beliefs concerning off-the-equilibrium path actions are consistent with each such action being taken by the type with the smallest cost of choosing that action. This assumption concerning off-the-equilibrium path actions is similar to the notion of a Proper Equilibrium first discussed in Myerson (1978). Also, in our model there are multiple equilibria that are identical except for how compensation is divided between salary and bonus because of differences in y^M. We focus on the equilibrium where the lower bound on the bonus always equals zero since this best matches the datasets we analyse in our empirical analysis.[16]

[16] Period 2 compensation contracts are individual specific so y^M is set such that the bonus equals zero if zero effort is chosen by the individual. For period 1 there is a single compensation contract for each group, so y^M for group k is set so the bonus equals zero if innate ability equals $\theta_k{}^L$ and effort equals zero. Note also that our assumption that the lower bound on the bonus equals zero could be derived endogenously by assuming bonuses cannot be negative and imposing Trembling Hand Perfection. The logic here is as follows. In period 1 a firm would not offer a contract with a higher value for y^M (which would also mean a higher salary) because the firm would lose money if the worker's ability was close to $\theta_k{}^L$ and the worker mistakenly chose zero effort. Similarly, a group k worker would prefer our contract with a lower bound of zero for the bonus over a contract with a lower value for y^M (which means a lower salary) because the contract with the lower value for y^M would yield lower worker utility if the worker subsequently mistakenly chose zero effort. Similar arguments yield that the lower bound for the bonus in period 2 must also equal zero given Trembling Hand Perfection.

As a final theoretical point, for tractability and expositional reasons we have conducted our theoretical analysis in a two-period setting and one might wonder the extent to which results generalize to a setting with more periods. We have conducted a preliminary analysis of a related three-period setting and our main results seem robust to moving to this three-period setting.[17] Specifically, our preliminary analysis indicates that results analogous to those found later in Corollaries 1 through 5 hold in this three-period setting which suggests this three-period model is consistent with all of our testable implications. We have chosen to focus on a two-period model in our formal analysis because the logic of the analysis is easier to follow in the two-period setting.

2.2 Preliminary Analysis

Because one of our goals is deriving how the bonus varies with job level holding age fixed, we restrict the analysis to parameterizations such that group 1 workers are assigned to job 1 in period 1 and group 2 workers are assigned to job 2 in period 1. Also, to simplify the analysis we focus on parameterizations such that in period 2 a worker is either kept at the same job level or promoted one level, i.e., the probability of being promoted two job levels or being demoted both equal zero. And we assume parameters are such that the probabilities of promotion are below the efficient levels which is the standard case in promotion signalling models. See Appendix A for details.

In the rest of this section we analyse how the model works period by period, where we start with period 2 and then consider period 1. In the next section we describe the full equilibrium of the model and discuss testable implications.

[17] In the three-period model there are three worker groups, where our focus as in the model analyzed in the text is parameterizations such that all group k workers, for each k, k=1,2,3, are on job k in period 1 and in each subsequent period a worker either remains at the same job level as in the previous period or moves up one level.

As indicated, we begin with period 2. Because of firm specific human capital, there is no turnover.[18] The contracting problems each firm faces in period 2 are standard contracting problems in which a firm chooses a salary, a minimum output level, and a bonus rate to maximize period 2 profits subject to participation and incentive compatibility constraints. Lemma 1 characterizes the solution to these maximization problems. Below let e_{ijk2} be the equilibrium effort choice of worker i in group k assigned to job j in period 2, α_{ijk2} be the worker's salary, y_{ijk2}^M be the minimum output specified in the worker's contract, β_{ijk2} be the bonus rate, z_{ijk2} be the training cost, and y_{ijk2} be the worker's output. Also, let θ_{ijk2} be the innate ability level of worker i in group k assigned to job j in period 2 and $U^{2M}(\theta_{ijk2})$ be the second period utility that worker i in group k assigned to job j by the worker's first period employer would receive by moving to a new firm at the beginning of period 2.[19]

LEMMA 1: *Equilibrium period 2 compensation contracts and effort levels satisfy i) through iii).*

i) $\beta_{ijk2}=1$ *and* $e_{ijk2}=e_j^*(s_2)$ *for every worker i in group k, k=1,2, assigned to job k in period 2 at the worker's first period employer, i.e., j=k.*

ii) $\beta_{ijk2}=1$ *and* $e_{ijk2}=e_j^*(s_1)$ *for every worker i in group k, k=1,2, assigned to job k+1 in period 2 at the worker's first period employer, i.e., j=k+1.*

iii) $y_{ijk2}^M=s_{k-j+2}(c_j+d_j\theta_{ijk2})-z_{ijk2}$ *and* $\alpha_{ijk2}+(y_{ijk2}-y_{ijk2}^M)-g(e_{ijk2})=U^{2M}(\theta_{ijk2})$ *for every worker i in group k assigned to job j in period 2, k=1,2 and j=k,k+1.*

---

[18] If we introduced turnover, then we could derive predictions concerning how bonus size varies with firm tenure. We do not take this step because the single firm dataset that we employ in our empirical analysis has noisy data concerning firm tenure due to the dataset only including the managerial part of the workforce.

[19] Consistent with the discussion at the end of the description of the model in Section 2.1, second period behavior described in Lemma 1 is unchanged with various different assumptions concerning the timing of contract offers for a worker at the beginning of period 2. The reason is that, given the form of compensation contracts assumed, competition for a worker between prospective employers translates into the worker's utility associated with moving being equal to the worker's productivity at an alternative employer given efficient effort minus the disutility associated with that effort level. In turn, the first period employer offers a contract such that the worker's utility associated with staying just matches this utility of moving whether the first period employer moves last through a counter-offer or offers are simultaneous.

Lemma 1 tells us that an equilibrium compensation contract in period 2 works quite simply. The contract induces the efficient effort level where the efficient effort level equates the marginal benefit of increased effort with the marginal cost of effort. This is captured in i) and ii), where in period 2 the firm always equates the marginal benefit of increased effort with the marginal cost of effort by setting the bonus rate equal to one. Note that, since period 2 is the last period, there are no promotion incentives driving effort choices in period 2, so the bonus rate is set such that a worker receives the full extra productivity associated with an increase in effort.

The other aspect of Lemma 1 is the determination of period 2 salaries and minimum output levels captured in iii). As mentioned earlier, we assume there is firm specific human capital so there is no turnover in this model and we also assume that the lower bound on the bonus equals zero. This means two things in equilibrium. First, a worker's minimum output level is what the worker produces given zero effort. Second, the first period employer chooses a compensation contract for each worker such that the worker's utility associated with staying just equals the utility associated with the worker moving to a new firm.

Now consider period 1. At the beginning of period 1 workers within each group look identical and so there are two compensation contracts offered to first period workers – a group 1 contract and a group 2 contract. In particular, the equilibrium first period compensation contracts are the ones that maximize expected worker utility over the two periods subject to an incentive compatibility constraint and a non-negative expected profit constraint. The logic is that competition between firms for workers at the beginning of the first period results in the equilibrium contracts being the ones that maximize expected worker utility subject to firms not losing money. Further, in equilibrium the non-negative expected profit constraint is binding so the equilibrium contract for each worker group is the one that maximizes expected worker utility subject to a zero expected profit constraint.

Let e_{ijk1} be the equilibrium effort choice of worker i in group k assigned to job j in period 1, α_{ijk1} be the worker's salary, y_{ijk1}^M be the minimum output specified in the worker's contract, β_{ijk1} be the bonus rate, and y_{ijk1} be the worker's output. Remember, the analysis is restricted to parameterizations such that group 1 workers are assigned to job 1 in period 1 and group 2 workers are assigned to job 2. Lemma 2 characterizes the solutions to these first period maximization problems.

LEMMA 2: *Equilibrium period 1 compensation contracts satisfy i) and ii).*

i) $\beta_{ijk1}=\beta_{jk1}<1$ *and* $e_{ijk1}=e_j^*(1)$ *for every worker i in group k, j=1,2, and j=k.*

ii) $\alpha_{ijk1}=\alpha_{jk1}$ *and* $y_{ijk1}{}^M=y_{jk1}{}^M$ *for every worker i in group k, j=1,2, and j=k, where* $\alpha_{jk1}+\beta_{jk1}[c_j+d_j(E(\theta^k)+e_j^*(1))-z-y_{jk1}{}^M]>c_j+d_j(E(\theta^k)+e_j^*(1))-z.$

Part i) states that effort choices in period 1 are efficient just like in period 2. The difference is that, as also captured in i), firms do not set the bonus rate for a worker in group k such that the marginal increase in the bonus due to increased effort just equals the marginal increase in productivity due to increased effort. Rather, bonus rates are set so that the marginal increases in the bonus due to increased effort are below the marginal increases in productivity. The reason is that workers perceive an increased probability of subsequent promotion from an increase in effort (we elaborate on this point in the next section) and firms take this into account in setting period 1 bonus rates.

The other result captured in Lemma 2 concerns how salaries and minimum outputs are determined. As captured in Lemma 1, because of firm specific human capital and asymmetric learning, a firm earns positive expected period 2 profits from hiring a worker in period 1. Since competition in period 1 means firms earn zero expected profits over the two periods from hiring a worker in period 1 and bonus rates are chosen in the manner that achieves efficient effort choices, it is the choice of salaries and minimum output levels that achieves this zero profit condition. The end result, as captured in the lemma, is that expected compensation for each worker in period 1 exceeds the worker's expected period 1 output.

3. Equilibrium and Testable Implications

In this section we describe the full equilibrium of the model which follows from the preliminary results in the previous section. We then derive testable implications. As was true for the preliminary results in the previous section, throughout we impose a set of parameter restrictions that ensure that a group k worker is assigned to job k in period 1, a group k worker is assigned to either job level k or job level k+1 in period 2, and probabilities of promotion are strictly below efficient levels as is standard in promotion signalling models.

Let θ_k' be the critical value for innate ability for old workers in group k such that in period 2 it is efficient to assign old worker i in group k with previous experience at the current employer to job k+1 when $\theta_i>\theta_k'$ and to job k when $\theta_i<\theta_k'$. To be precise, θ_k' satisfies $s_2[c_k+d_k(\theta_k'+e_k^*(s_2))]-g(e_k^*(s_2))=s_1[c_{k+1}+d_{k+1}(\theta_k'+e_{k+1}^*(s_1))]-g(e_{k+1}^*(s_1))-z$. Also, we assume parameters are such that $\theta_1^L<\theta_1'<\theta_1^H<\theta_2'$ and $\theta_1'<\theta_2^L<\theta_2'<\theta_2^H$. That is, it is efficient in period 2 for a θ_k^L worker to be assigned to job level k by the first period employer and for a θ_k^H worker to be assigned to job level k+1.

Proposition 1 describes equilibrium behaviour in our model. Also, below a promotion refers to a worker being assigned to a job level one level higher than the worker was assigned to in the previous period, while not being promoted means the worker is assigned to the same job level as in the previous period.

PROPOSITION 1*: There exist values θ_1^+ and θ_2^+, $\theta_1^+>\theta_1'$ and $\theta_2^+>\theta_2'$, such that equilibrium behaviour is described by i) through iii).*

i) *In period 1 every worker in group k, k=1,2, is assigned to job k by the worker's period 1 employer and compensation contracts and effort choices satisfy i) and ii) of Lemma 2.*

ii) *In period 2 every worker i in group k, k=1,2, is promoted by the first period employer if $\theta_i\geq\theta_k^+$ and not promoted if $\theta_i<\theta_k^+$.*

iii) *In period 2 every worker i in group k, k=1,2, stays with the first period employer and the compensation contracts and effort choices satisfy i), ii), and iii) of Lemma 1.*

The proposition combines results from the previous section and also introduces the new result that, as is standard in promotion signalling models, there is a distortion in the promotion decision, i.e., $\theta_k^+ > \theta_k'$ for k=1,2. The logic for the distortion is the same as found in many earlier papers in the literature. When a worker is promoted at the beginning of period 2 by the worker's first period employer a positive signal is sent to prospective employers about the worker's ability. As a result, a promotion causes the compensation associated with moving to increase, which in turn means that in order to retain a promoted worker the first period employer must also make its compensation contract more attractive. Since making the compensation contract more attractive is costly to the first period employer, there is a promotion distortion in the sense that firms do not promote workers who are only slightly more productive on the higher level job.

We now provide a brief discussion of the role of the training cost in equilibrium behaviour in our analysis. As analysed and discussed in Waldman and Zax (2016), models of promotion signalling do not always result in promotion compensation increases due to the signal and when there is no promotion compensation increase due to signalling there is also no promotion distortion. For example, in Golan's (2005) promotion signalling model which adds a counter-offer assumption to Waldman's (1984a) model, promotions serve as a positive signal of worker ability but there is no promotion wage increase due to signalling and promotion decisions are efficient. The logic follows. In that model, because of the counter-offer assumption, there is a winner's curse result which means the poaching wage equals the lowest productivity associated with the signal. If worker ability does not affect output on the low level job as Golan assumes following Waldman (1984a), then efficient promotion decisions are such that the lowest ability worker promoted has the same output in the high level job as the worst worker in the low level job. As a result, even though a promotion signals high ability in equilibrium, the poaching wage does not rise with a promotion if promotion decisions are efficient and promotion decisions as a consequence are indeed efficient.[20]

In our model, if there is no training cost, then results are related to those found in Golan (2005). Specifically, promotions would serve as a positive signal of worker ability but for the marginal worker promoted the promotion would not necessarily result in a compensation increase. And if there is no compensation increase, then there would be no promotion signalling distortion. The logic is related to the logic above for Golan's (2005) results. Without a training cost, efficiency in promotion decisions means the marginal worker promoted does not necessarily receive higher utility due to the promotion if she chooses to switch employers and when this is the case compensation paid by the current employer does not rise given a promotion. Given this, the current employer would not distort the promotion decision by having too few promotions. With the training cost included and given it is sufficiently large, however, efficient promotion decisions do result in the marginal worker promoted receiving higher utility due to the promotion if she chooses to switch employers. In turn, this causes the current employer to distort promotion decisions in a fashion similar to that found in Waldman (1984a) and elsewhere.[21]

A further point to note is that, although in Proposition 1 promotions occur for each group of workers when innate ability is above some critical value, we could have instead written the proposition to state that a promotion occurs when the worker's first period output is above some critical value. In this model first period employers observe first period outputs and correctly infer workers' innate ability levels. So, saying that a worker's innate ability level is above some critical value and is promoted is equivalent to saying that the worker's first period output was above some critical value and this is what led to the promotion. And it is this relationship

[20] Waldman and Zax (2016) show that employing Golan's set-up but allowing worker ability to have a positive effect on worker productivity in the low level job results in inefficient promotion decisions similar to findings in Waldman's original analysis and elsewhere.

[21] Without a training cost our results would be qualitatively the same if we assumed that task specific human capital is mostly firm specific, i.e., h_2-h_1 is small relative to s_2-s_1.

between first period output and second period promotion which causes the possibility of subsequent promotion to serve as an incentive for first period effort.

Overall, the main point of the proposition is that in this model incentives stem from two sources: current monetary payments for high output due to the bonus and in period 1 future monetary rewards due to the promotions that follow when a worker produces high output. Further, the bonus rate is always set so that a worker chooses the efficient level of effort.

We now turn to testable implications, where our focus is variation in the size of the bonus payment. We start with results concerning how bonus payments vary with job level.

COROLLARY 1: *Holding job level tenure constant, the average bonus payment strictly increases with job level for old workers and also, if δ is sufficiently small, strictly increases with job level for young workers.*

Corollary 1 says that the average size of bonus payments rises with job level holding job level tenure and worker age fixed. There are two main reasons for this. First, because the marginal increase in output with respect to effort rises with job level, workers at higher job levels choose higher effort levels which translate into higher bonuses. Second, the higher marginal increases in output with respect to effort at higher job levels directly translate into higher bonuses.

One question is what is the role of the condition that δ must be small for this result to hold unambiguously for young workers. The answer concerns the trade-off in our model between bonus incentives and promotion incentives. In our model for young workers bonus incentives are set to augment promotion incentives so that effort choices are efficient. It is possible that promotion incentives rise so much at higher levels of the job ladder that bonus incentives fall and, in particular, the average bonus falls. If δ is small, then the difference in promotion incentives between young workers assigned to levels 1 and 2 is second order.

Note that, if δ is not assumed to be small, then consistent with an earlier discussion it is possible that promotion incentives could be very large thus limiting the use of bonuses. For example, in our model if δ is large and both d_2-d_1 and d_3-d_2 are also large, then the incentive for being promoted to job levels 2 and 3 could be high enough that the average bonus payment on levels 1 and 2 for young workers would both be very small. Potentially with a dataset with detailed information about tasks and productivity at each job this prediction could be tested. A related idea appears in DeVaro *et al*. (2018).

The next result concerns the relationship between bonus size and job level tenure.

COROLLARY 2: *For old workers on job level 2, holding all other parameters fixed, the average bonus payment increases with job level tenure.*

The first thing to note about Corollary 2 is that it only refers to old workers on job level 2. The reason is that in our model this is the only group of workers for whom job level tenure in fact varies. For example, all young workers on either job level are in their first period at that level, while old workers on level 1 are all in their second period at that level. The basic logic behind the corollary is that task specific human capital increases productivity for workers with previous experience at that level and the higher productivity increases equilibrium effort which in turn increases the bonus payment.

In Corollary 3 we consider the relationship between bonus size and worker age.

COROLLARY 3: *For workers on job level 2, holding job level tenure constant, the average bonus payment increases with worker age if s_1 is sufficiently large.*

Corollary 3 refers only to workers on job level 2 because that is the only job level for which workers assigned to that level vary in terms of age after controlling for job level tenure. The corollary states that bonus payments for these workers increase with age if human capital accumulation as workers age is sufficiently large. There are two reasons that bonus payments in this model increase with age after controlling for job level and job level tenure. First, old workers choose higher effort levels because of human capital accumulation. Second, old workers have no promotion incentives since this is their last period in the labour market, so achieving efficient effort incentives requires a higher bonus. But there is a third factor. Remember that we assume that the lower bound on the bonus equals zero. This means that for young workers the expected bonus includes the bonus rate multiplied by the average innate ability in a group minus the minimum innate ability in that group while old worker bonuses do not include such a component. The condition that s_1 is sufficiently large means this factor is dominated by the other two which, in turn, yields the prediction that bonus payments should rise with age.

We now consider how performance is related to the size of bonus payments.

COROLLARY 4: *Holding job level fixed, bonus payments rise with output for young workers.*

This prediction is not surprising. It simply says that bonus payments increase with performance which is true for young workers as long as the bonus rate is positive. Given our assumption that δ is sufficiently small that the bonus rate is always positive, we have that bonus payments increase with performance for young workers. For old workers there is no variability concerning the bonus payment, holding job level and job level tenure fixed, because there is no stochastic term in the production functions. If we introduced a stochastic term, then there would also be a positive relationship between performance and the size of bonus payments for old workers.

Our last prediction concerns how changes in promotion based incentives affect bonus size. In order to explore this relationship, we conduct a comparative statics analysis of how bonus size is affected by a change in δ. The basic logic is that as δ increases, i.e., there is less discounting, promotion incentives become more important in a worker's choice of effort in period 1.

Corollary 5 formally states what happens to bonus size for young workers when δ increases.

COROLLARY 5: *An increase in δ, holding all other parameters fixed, results in a decrease in the average bonus payment for young workers on each job level.*

The logic is again captured by the idea that in our model the bonus is always set equal to the value that achieves the efficient effort level. Increasing δ increases promotion incentives for young workers. But an increase in δ has no effect on the efficient effort levels for young workers. So, when promotion incentives rise but there is no change in the efficient effort levels, the result is a decrease in the size of bonuses required to achieve efficient effort levels.

Note that the more general theoretical prediction here is that an increase (decrease) in the promotion prize results in lower (higher) bonus payments. In the corollary we model this as a change in the value of the promotion prize due to a change in discounting but any change which affected the value of the promotion prize but not the efficient effort level on the low level job would result in the same prediction. See DeVaro and Waldman (2012) and Bognanno and Melero (2016) for analyses that show that in promotion signalling models there are various factors that affect the size of promotion prizes that are independent of the efficient effort levels at the low level jobs.

In summary, our model has five testable implications. First, bonus size should rise with job level holding job level tenure and worker age constant. Second, bonus size should rise with job level tenure holding both job level and age constant. Third, bonus size should rise with age holding job level and job level tenure constant. Fourth, bonus size should rise with performance holding job level and age fixed. Fifth, an increase in promotion incentives which we model as an increase in δ should cause bonus size to decrease. Note that one of the predictions depends on human capital accumulation being substantial but this is likely the case in the datasets we focus on in our empirical analysis for significant portions of workers' careers given the nature of age-earnings profiles. We come back to this issue later when we discuss our empirical findings concerning bonuses and age.

4. Data

We use two different datasets to test our theoretical predictions. The data used in our first set of tests comes from the personnel records of a medium-sized US firm operating in the financial services industry. This dataset was first analysed in the seminal studies of Baker *et al*. (1994a,b) that focused on various aspects of the internal labour market operations of this firm. The full dataset covers all the managerial employees at the firm over the period 1969-1988 and includes salary, bonus, and subjective performance variables, as well as demographic variables including age, race, gender, and years of education. For our purposes, the variables of special interest are job levels, bonuses, and performance ratings.

Since the HR department at the firm did not provide any information about job levels, Baker, Gibbs, and Holmstrom constructed job level data from the raw data by using typical movements between job titles. In their original study, Baker, Gibbs, and Holmstrom identified eight levels, where level eight is the top level filled by the CEO. Since the manner in which CEO compensation is determined is likely different than the way compensation is determined for other firm employees and since there are few employees at the top levels, we drop observations from level five and higher. Subjective performance ratings are measured on a five-point scale, where 1 represents the best performance and 5 the worst. Note that performance ratings are not available for all observations in the dataset (72% of the sample we employ include a performance rating), so the sample is smaller when we include performance ratings in the regression specification. Also, because there are few observations with ratings equal to 4 or 5, we pool together all ratings equal to 3, 4, and 5.

Salaries and bonuses are reported annually and are measured in real terms in 1988 dollars. Bonuses are only reported for the time period 1981-1988, so that is the part of the sample we use (we do, however, use observations from earlier years to construct lagged values of some variables). Bonuses for a given year are paid in February of the following year, where not all eligible employees earn a bonus (about 34% of all worker-years in our sample include a strictly positive bonus).

In addition to restricting our sample to the time period in which bonus data is available, we also restrict our sample in two other ways. First, since compensation data are in local currencies, we only include observations of workers employed in US plants. Second, we drop observations in which the worker received a demotion (there are only 47 observations that include demotions in the time period we study).

These restrictions leave us with a sample consisting of 23,637 worker-years and 7,498 workers. Summary statistics are reported in Table 1. The average worker is 39 years old and the average value for tenure at the firm is 5.6 years. Focusing on job level, we see that as job level increases workers on average are older, have higher values for tenure, receive better performance ratings, and earn larger bonuses. It is also interesting to note that the bonus-salary ratio increases with job level, i.e., the proportion of total compensation which comes from the bonus is higher at higher levels of the job ladder.

The data for our second set of tests are drawn from a large linked employer-employee panel dataset of Finnish firms. The source of this dataset is the records of the Confederation of Finnish Industries which is the central organization of employer associations in Finland.[22] The data are of high quality given that the data are based on firms' administrative records. Also, since participation in the survey is compulsory except for the smallest firms, the response rate is nearly 100%.

---

[22] See Kauhanen and Napari (2012) for a more detailed description of the data and of the wage-setting process in Finland.

We focus on white-collar employees in manufacturing because for these employees we have the necessary data to test most of our theoretical predictions. Importantly, in this sector all firms use the same job classification scheme which makes it possible to define the job ladder and promotions similarly at all firms. The job classification scheme includes 56 job titles and four job levels, where we define a promotion as a move to a higher job level. The data also includes information on salary, bonuses, and demographic variables such as age, gender, and education level. Bonuses for a given year are typically paid in the following year, so in our tests we use the following year's bonus data to test our theoretical predictions. The main difference with the Baker, Gibbs, and Holmstrom dataset is that the Finnish dataset does not include subjective performance evaluations. On the firm side there is information on industry and firm size can be calculated since the data covers practically all employees at each firm.

We restrict the analysis to the time period 2003 to 2012 because there was a change in the classification of job titles the previous year. Our analysis focuses on firm stayers, on job level/firm/year-cells in which bonuses are used, and firms that employ more than 1000 workers. These restrictions make the sample comparable to the Baker, Gibbs, and Holmstrom dataset. We also drop employees at the highest job level because by definition they cannot be promoted. Our sample consists of 387,687 person-years and 80,479 individual persons.

Summary statistics for this dataset are reported in Table 2. It is of interest to note that in terms of the main focus of the paper which is the determinants of variation in bonus payments this dataset is qualitatively similar to the Baker, Gibbs, and Holmstrom dataset. That is, bonus payments rise with job level and the bonus/salary ratio also rises with job level. There are, however, a few differences between the datasets. In the Baker, Gibbs, and Holmstrom dataset average age rises with job level which is not true here. Also, both average tenure at level and average tenure at firm rise with job level in the Baker, Gibbs, and Holmstrom dataset but neither is true in this dataset.

5. Empirical Analysis

The empirical analysis consists of three parts. We start with a preliminary examination of the firm's bonus policy. In particular, we provide an analysis of which worker attributes are related to the probability a bonus is received. We then test predictions concerning the relationships between job level, job level tenure, age, performance, and bonus payments. In our final set of tests, we focus on the prediction concerning the trade-off between bonus payments and incentives provided through promotions.

5.1 Some Preliminary Tests

In this subsection we focus on worker attributes correlated with the probability of receiving a bonus. In particular, we test the extent to which this probability is correlated with job level, job level tenure, performance, wage growth, average salary increase at the current level, and whether or not the worker was promoted. In our theoretical analysis basically everyone receives a bonus. However, if we added a small cost associated with paying a bonus, then small bonuses would not be paid in equilibrium which means the probability a worker receives a bonus would be positively related to worker effort which itself should be positively correlated with variables such as job level, job level tenure, and performance. Identifying the correlations between these variables and probability of receiving a bonus thus provides evidence concerning whether our general theoretical approach to modelling bonuses is correct. In the following subsections we focus on the specific testable predictions derived in the theoretical analysis.[23]

We estimate linear probability models where the dependent variable is an indicator variable that takes on a value of one if the worker received a bonus in the given year and zero if not. Results for the single firm dataset are reported in Table 3. In columns 1 to 5 we do not include fixed effects while in columns 6 through 10 we add individual fixed effects. We start our discussion with columns 1 through 5. The explanatory variables in column 1 are indicator

---

[23] We do not refer to any tests in this subsection as tests of the theory since, as just discussed, taken literally our model predicts a bonus should basically always be paid.

variables for job level (the omitted category is level one) and controls for the worker's age, race, gender, and education as well as indicator variables for year. We find that the probability of a bonus payment increases with job level, where all coefficients are statistically significant at the 1% level and differences between coefficients are all statistically significant at the 1% level.

In columns 2 through 5 we add an indicator variable for whether the worker received a promotion in the following year, a variable capturing average salary increase at the current level, indicator variables for tenure at the current level (where the omitted category is tenure equals one, i.e., this is the worker's first year at the current level), and performance ratings. Adding these additional explanatory variables does not change the qualitative nature of our finding that the probability of receiving a bonus rises with job level. Further, we find that promotion and average salary increase at the current job level are both positively correlated with the probability of receiving a bonus, while this probability increases with performance (remember that a higher performance rating represents worse performance). Also, this probability is negatively correlated with tenure at current level when performance ratings are not included but this correlation mostly disappears in column 5 when performance ratings are added.

As indicated, in columns 6 through 10 we add individual fixed effects. The results are mostly qualitatively unchanged. The only clear difference is that now there is some evidence that the probability of receiving a bonus rises when tenure at job level increases from one to two and then three. However, this evidence is weak since the relevant coefficients mostly lose statistical significance when performance ratings are included.

Overall, the results in Table 3 provide a number of interesting insights concerning the single-firm dataset. For example, the probability of a bonus payment rises with job level as found in a number of earlier studies and the finding remains after controlling for various factors such as performance ratings and job level tenure. So, for example, the job level result is not driven by the possibility that the probability of a positive bonus payment is positively correlated with performance and average performance rises with job level. Also, not surprisingly, the probability of a bonus rises with performance which is consistent with our theoretical approach (and many others) which predicts bonuses rise in response to a high level of performance.

There is also a result suggestive of our theoretical prediction of a trade-off between bonus payments and incentives derived from the possibility of future promotion. One could imagine that, for many promotions, prior to the promotion the worker may have achieved a level of performance such that a promotion is warranted but promotion is delayed for a year or two (this does not arise in our model, but could arise if we added slot constraints). So, if bonus incentives and promotion incentives are substitutes, one might predict that bonus incentives would rise in periods in which performance has increased high enough to warrant a promotion but promotion is delayed. The findings in columns 2 through 5 and 7 through 10 are consistent with this prediction. That is, in each column the coefficient on the indicator variable for promotion in the following year is positive and statistically significant which is consistent with bonus incentives being higher when promotion is warranted but delayed.

In Table 4 we conduct the same tests as in Table 3 except we employ the Finnish multi-firm dataset. Also, because as mentioned earlier this dataset does not include performance evaluations, Table 4 does not include regressions analogous to those found in columns 5 and 10 of Table 3. Most of the findings in Table 4 are qualitatively similar to those found in Table 3 but there are a couple of differences. First, there is no evidence of workers on level two having a higher probability of receiving a bonus than workers on level one. In fact, without fixed effects in columns 1 through 4 the results indicate a lower probability of individuals receiving a bonus on level two than on level one, but with fixed effects the results indicate no difference between the two levels concerning the probability of receiving a bonus. Second, there is more evidence that a higher value for job level tenure translates into a higher probability of receiving a bonus. Without fixed effects all the coefficients on the tenure at level variables are positive and statistically significant at the 1% level indicating that the probability of receiving a bonus is lower in the first year at the level. But there is no indication that this probability goes up after the second year a worker is on a level. In the fixed effects specification, however, the coefficients on the tenure at level variables rise with tenure at level up to tenure at level equal to four and the differences are all statistically significant at least at the 5% level.

5.2 Job Level, Job Level Tenure, Age, and Bonus Payments

We now turn to the testable predictions derived in the theory section. In this subsection and the next we focus on the first four testable predictions which concern how bonus payments vary with job level, job level tenure, age, and performance. In the last subsection of this section we consider the fifth testable prediction which concerns the trade-off between bonus incentives and promotion incentives.

Before proceeding to the formal tests, it is useful to consider the basic data on bonuses categorized by job level. As seen in Tables 1 and 2, the average bonus payment increases with job level, where the overall bonus structure is convex in the sense that the increase between adjacent levels is higher at higher levels. It is also the case that the bonus/salary ratio rises with job level. Our theory predicts that the positive relationship between bonus size and job level should hold even after controlling for job level tenure, worker age, and other control variables, while there should also be a positive relationship between bonus size and job level tenure after including controls and bonus size and age after including controls. This is what we consider next.

To empirically test these predictions we employ the following econometric specification.

$$\beta_{ijt} = Z_i\varphi + X_{it}\tau + \alpha_{it}\gamma + \sum_{j=2}^{J} L_{it}^{j}\delta^{j} + \Psi_{it}^{e}\rho + t_i + \mu_i + \varepsilon_{it} \qquad (2)$$

In equation (2) β_{ijt} is the bonus payment made to worker i on job level j in year t; Z_i is a vector of time-invariant attributes of worker i which includes indicator variables for the worker's race, gender, and education level; X_{it} is a vector of time-varying attributes of worker i which includes tenure at the current job level and performance ratings at year t to capture variation both across workers and for a given worker over time in productivity; α_{it} is a vector that includes the age of worker i in year t and its squared term (divided by 100 for convenience); L_{it}^{j} is a level-specific binary indicator variable, where $L_{it}^{j}=1$ if $L_{it}=j$ and 0 otherwise and L_{it} is the job level of worker i in year t;[24] Ψ_{it}^{e} is an estimate of the size of the expected promotion prize for worker i in period t; t_i is a vector of year indicator variables used to control for the effect of the business cycle on bonus payments; μ_i is a worker-specific unobserved factor that may be correlated with other explanatory variables;[25] and ε_{it} is an idiosyncratic error term that is independently and identically distributed with mean zero. The variable Ψ_{it}^{e} measures the expected future compensation increase for worker i in period t associated with the possibility that worker i is promoted in period t+1. The final subsection of this section focuses on the prediction that bonus payments should be negatively related to the size of promotion incentives. In that subsection we describe how we estimate for each observation the expected promotion prize variable.

A crucial point in testing for the effect of job level on bonus payments is the assignment of workers to job levels. As our theoretical model illustrates, more able workers are assigned to higher job levels in equilibrium. Even though proxies for performance and ability are included in the regressions, a part of the variation that affects job assignment is likely not captured by these variables. As a result, indicator variables for job levels may be correlated with the disturbance term and their point estimates may consequently be biased. With this in mind, we begin the analysis with the ordinary least squares (OLS) estimation on pooled data. Since OLS estimation requires the most rigid conditions to produce unbiased estimates, results from the pooled regressions are used as a benchmark. Then, in order to mitigate the effect of unobserved worker heterogeneity, we make use of the panel dimension of the data by employing a fixed effects estimation which relaxes conditions required by the OLS estimation.[26,27]

---

[24] Level one is the omitted category.

[25] We do not employ a random-effects estimation in the empirical analysis since its restriction that the worker-specific unobserved factor must be uncorrelated with the explanatory variables is not realistic in the current model.

[26] In our theoretical model, many of our predictions hold both for a given worker and on average. But changes in the model such as having output be a function of ability times effort rather than ability plus effort would result in worker ability and bonus size being correlated. Given this, considering regressions with and without fixed effects seems like the best approach.

[27] Technically, the OLS estimation yields unbiased estimates if ability differences that affect worker assignments to job levels are fully accounted for by the variables used in the specification, i.e., $E[L_{it}^{j} \cdot \eta_{it} \mid Z_i, X_{it}, \alpha_{it}, t_i]=0$ for all t and j, where $\eta_{it}=\mu_i+\varepsilon_{it}$. The fixed effects estimation, on the other hand, relaxes this condition by assuming that the

We again begin with the Baker, Gibbs, and Holmstrom dataset. Table 5 shows the results of estimating equation (2) for alternative specifications. Columns 1 and 2 show the OLS pooled regression results with and without performance ratings included, while columns 3 and 4 show the fixed effects regression results with and without performance ratings included. Controls for gender, race, education level, and year indicator variables are included in each of the pooled regressions, while they are dropped in the fixed effects regressions. Note that the number of observations in columns 2 and 4 is smaller because of missing performance ratings for some observations. Recall that the first prediction of our theoretical model is that, holding job level tenure and age constant, bonus payments increase with job level, i.e., $\delta^{j+1}>\delta^{j}>0$ for j=1,2, and 3.

The results reported in the first four columns of Table 5 provide clear support for our first theoretical prediction. In each regression all three coefficients on the job level indicator variables are positive as predicted by the theory and also statistically significant at the 1% level. Also, in seven of the eight cases the coefficient on the job level indicator variable is larger than the coefficient on the indicator variable for the lower adjacent level, where most of these differences are themselves statistically significant at the 1% level. The results also suggest that unobserved worker heterogeneity plays an important role in workers' assignments to job levels. This follows since most of the coefficients on the job level indicator variables rise sharply when we include worker fixed effects.

Our second testable prediction is that bonus size should vary positively with job level tenure holding job level and age fixed. The first four columns of Table 5 provide support for this prediction, especially at lower levels of job tenure. Let us start with the pooled OLS regression results in columns 1 and 2. In column 1 which does not include performance ratings the coefficient on each tenure at level variable is positive and statistically significant at either the one or five percent level which is consistent with the theoretical prediction. Also, the coefficients rise with job level tenure except for the coefficient on the tenure at level equal to five variable

---

unobserved attributes of workers that affect their job assignment are fully captured by the time-invariant individual-specific factor, i.e., it requires that $E[L_{it}^{j} \cdot \varepsilon_{it} \mid Z_i, X_{it}, \alpha_{it}, \mu_i]=0$ for all t and j which is smaller than the coefficients for tenure at level equal to three and four. It is the case, however, that the differences between coefficients at adjacent tenure at level variables are mostly not statistically significant.[28] In column 2 we add performance ratings and the main difference is that the coefficient on the tenure at level variable only rises up to tenure at level equal to three.

Now consider columns 3 and 4 which include fixed effects. There are a couple of differences when fixed effects are added. First, coefficients are larger in the fixed effects regressions which again suggests unobserved worker heterogeneity plays an important role in workers' assignments to job levels. Second, statistical significance is weaker for the coefficients for the tenure at level greater than or equal to six variable and that was not the case for the OLS regressions in columns 1 and 2.

One interesting finding concerns the effect of adding performance ratings on the coefficients concerning the tenure at level variables. These coefficients are predicted to be positive and rising with tenure at level because performance rises with tenure at level due to task specific human capital. This suggests that adding performance ratings should decrease the sizes of the coefficients on the tenure at level variables. This pattern arises in the fixed effects regressions in columns 3 and 4 but not in the OLS regressions in columns 1 and 2.

Our third testable prediction is that bonus payments should increase with worker age holding job level and job level tenure fixed. The results in columns 1 through 4 of Table 5 do not support this prediction. In the pooled OLS regression results in columns 1 and 2 there is only one statistically significant coefficient on the age variables and it does not indicate a positive relationship between bonus size and age. In the fixed effects regression results in columns 3 and 4 the only positive and statistically significant coefficients related to age are on the age squared variable. These coefficients are consistent with age being positively related to the size of bonus payments for old enough workers. But, as discussed further below, a nuanced understanding of

[28] But some of the differences in coefficients across tenure at level variables further apart are statistically significant. For example, the coefficients on the tenure at level variable greater than or equal to six is statistically significantly different than the coefficient on the tenure at level variable equal to two at the 5% level.

the theory suggests the positive correlation should hold for younger workers and not for older workers. We come back to a discussion of these results below.

The last theoretical prediction we consider in this subsection is that bonus size should increase with performance even after controlling for age, job level, and job level tenure (as indicated earlier, various other models of bonus payments would also make this prediction). The results in columns 2 and 4 of Table 5 provide clear support for this prediction. In the OLS regression in column 2 which includes performance ratings the coefficient on each performance rating variable is negative and statistically significant at the 1% level which is consistent with the theoretical prediction (remember, a higher performance rating means worse performance). Also consistent with the prediction, the coefficients fall with the performance rating and the difference between the two coefficients is statistically significant at the 1% level. Further, in column 4 which adds fixed effects basically the same pattern is repeated.

In the first two columns of Table 6 we reproduce using our multi-firm dataset the tests in the first four columns of Table 5 that do not include performance ratings. Notice this means that there are no regressions in Table 6 analogous to the regressions in columns 2 and 4 of Table 5. The first two columns of Table 6 are similar to the first four columns of Table 5 in that they provide clear support for the first two testable predictions. That is, in both regressions there is evidence that bonus payments rise with job level and also evidence that bonus payments rise with tenure at level at least for lower values of job level tenure, although in the fixed effects specification the evidence points to lower bonuses on level two than on level one. One other difference between the tables is that in Table 5 there is some evidence (see column 1) that bonus payments rise with tenure at level up through tenure at level being greater than or equal to six, while in Table 6 there is evidence that bonus payments rise with tenure at level up through tenure at level equal to five but the evidence indicates that it falls when tenure at level becomes greater than or equal to six.

The main difference between the first four columns of Tables 5 and the first two columns of 6 concerns the third testable prediction which is that bonus payments should rise with age holding job level and job level tenure fixed. As discussed above, the first four columns of Table 5 show no support for this prediction at lower ages. In contrast, in Table 6 the coefficients on the age variable in columns 1 and 2 are positive and statistically significant at the 1% level, while the coefficients on the age squared variable in these regressions are negative and statistically significant at the 1% level. These results suggest that, holding job level and job level tenure fixed, bonus payments rise with age for a significant portion of workers' careers. To be specific, given our specification, the effect of age on bonus payments is measured by the semi-elasticity of bonus payments with respect to age. From equation (2) we can derive that this elasticity term is given by $\gamma_1+y(1/50)\gamma_2$, where y is the age level at which the elasticity is evaluated. The bottom panel of Table 6 shows the F-statistic and the associated p-value for the null hypothesis that the semi-elasticity of bonus payments with respect to age is zero for the average worker. For the regressions reported in the first two columns these tests suggest that in this dataset age matters.

We can use this formula and the coefficients reported in the first two columns of Table 6 to derive estimates of the age range in our Finnish dataset for which, holding job level and job level tenure fixed, bonus size increases with age. The coefficients in column 1 indicate that bonus size increases with age up to age approximately equal to 43 while the column 2 coefficients yield that bonus size increases with age up to age approximately equal to 55.[29] In other words, both columns 1 and 2 yield that bonus size increases with age at younger ages but not at sufficiently high ages.[30]

At first one might think that our theoretical approach is inconsistent with the size of bonus payments falling with age at high ages. But we believe, in fact, that this result is

---

[29] The calculation for column 1 is $0.006-y(1/50)(0.007)=0$ which yields $y\approx43$. For column 2 the calculation is $0.012-y(1/50)(0.011)=0$ which yields $y\approx55$.

[30] We have also run our OLS and fixed effects regressions allowing for interactions between the age variables and job level variables and also between tenure at job level and job level. Introducing either type of interaction has no effect on our qualitative findings that bonus size increases with job level, increases with tenure at level, especially at low values for this variable, and decreases with the performance rating. In terms of the test that includes interactions of the age variables with the job level variables, the results do not change the basic conclusions of the tests without these interactions. That is, in the BGH dataset there is little evidence that bonus size rises with age, while in the Finnish dataset it rises with age at low ages. In terms of the test interacting tenure at level with job level, the fixed effects results suggest that the positive effect of job level tenure on bonus size may be smaller at higher job levels. These tests are available from the authors upon request.

consistent with our theoretical approach. Our theoretical approach predicts that bonus size should continue to increase with age as long as higher ages are associated with substantial increases in human capital. If, however, at high enough ages human capital increases slowly or decreases with further increases in age, then the predicted relationship between age and bonus size becomes ambiguous.

The reason this point is relevant is that studies of age-earnings profiles and experience-earning profiles suggest that human capital peaks at some point and, in fact, decreases with age at high ages. For example, using a quadratic specification, Murphy and Welch (1990) found that for workers with a high school education or higher (which is the case for a large proportion of our sample) the experience-earnings profile peaks at around twenty five years of labour market experience. Similarly, Asplund (2001) finds that for the Finnish workforce during the time period of our study experience-earnings profiles peak around 31 to 38 years of experience (the specific value depends on year and gender). If we interpret the Asplund results to mean that human capital also peaks at approximately that level of labour market experience, then our theoretical approach would seem to be consistent with the results concerning age found in columns 1 and 2 of Table 6.[31]

One question of interest is why do the two datasets provide such different results concerning age? One possibility concerns managerial stock holdings and stock options. The Baker, Gibbs, and Holmstrom dataset concerns a bank where managerial stock holdings and stock options are likely quite important, while the Finnish dataset concerns manufacturing firms where managerial stock holdings and stock options are quite rare (see Kauhanen and Napari (2012) for a discussion).[32] The reason that managerial stock holdings and stock options, if they are heavily used in a firm, can be important in terms of our prediction concerning age is that

---

[31] In our sample the average starting age is about 19. Combining this with the Asplund result that experience-earnings profiles peak between 31 and 38 years of experience suggests that bonus size should peak between 50 and 57. This matches the column 2 calculation quite well while the column 1 calculation seems a little low.
[32] Unfortunately, the Baker, Gibbs, and Holmstrom dataset does not include data on stock holdings or stock options so we cannot test this argument directly.

stock holdings and stock options are an additional avenue through which a firm can provide incentives. To see this, consider a firm in which increases in incentives achieved through stock ownership and stock options as managers age match the increase in non-promotion incentives needed to maintain efficient incentives as workers age. Our approach predicts that in such a firm there will be no correlation between bonus payments and age. We suspect that this possibility is important in the Baker, Gibbs, and Holmstrom dataset but not in the Finnish dataset which would explain why the age prediction only holds in the Finnish dataset.

5.3 Alternative Empirical Specifications

In our theoretical model we assume a compensation contact consisting of a base salary, α, a minimum output level, y^M, and a bonus rate, β. The bonus payment, b, is thus given by $b=\beta(y-y^M)$. Using the equilibrium compensation contracts for periods 1 and 2 as shown in Lemma 1 and Lemma 2, respectively, the equilibrium bonus payment is given in equation (3).

$$\begin{aligned} b_{ijkt} &= \beta_{ijkt}[y_{ijkt} - y_{ijkt}{}^M] \\ &= \beta_{ijkt}[s_{it}(c_j + d_j(\theta_i + e_{it})) - z_{ijt} - (s_{it}(c_j + d_j\theta_i) - z_{ijt})] \\ &= \beta_{ijkt} s_{it} d_j e_j{}^*, \end{aligned} \tag{3}$$

where $e_j{}^*$ denotes the efficient effort level and is determined by $s_{it}d_j=g'(e_j{}^*(s_{it}))$, for j=1,2,3. Notice that the bonus payment does not depend directly on worker ability but worker ability indirectly affects the bonus payment since it determines the worker's job assignment which, in turn, affects the bonus payment. Under the assumptions discussed in footnote 27, using fixed effects to estimate equation (2) will yield unbiased results.

Under alternative assumptions, however, the fixed effects specification may yield biased results. The following argument is related to discussions found in Gibbons *et al.* (2005) and Lluis (2005). Equation (2) can be rewritten in the simpler form given in equation (4).

$$\beta_{ijt} = X_{it}\tau + \sum_{j=2}^{J} L_{it}{}^j\delta^j + \mu_i + \varepsilon_{it} \tag{4}$$

To estimate the parameters τ and δ^j, j=2,3,4, consistently, we need to explain how we treat the worker's ability, μ_i. Under the assumption that μ_i is observed by firms but not by the econometrician, (4) can be estimated using fixed effects. It is easy to see that first differencing (4) eliminates worker fixed effects. This is the approach taken in the previous subsection.

As pointed out by Gibbons *et al.* (2005), there are two limitations associated with this approach. First, if firms learn about a worker's ability over time, then observed ability becomes time variant. Second, if the returns to worker ability differ across job levels, an interaction term between job level dummies and unobserved worker ability enters the estimating equation. In both cases, first differencing the estimating equation does not eliminate worker fixed effects.

We start with the first case. Suppose firms use past realizations of wages, bonuses, performance ratings, etc., to form a belief about a worker's ability. Let μ_{it}^e denote the firm's belief about worker i's ability at time t. Given Bayesian beliefs are a martingale, we have

$$\mu_{it-1}^e = \mu_{it-2}^e + \zeta_{it-1}, \tag{5}$$

where ζ_{it-1} is orthogonal to μ_{it-2}^e. Replacing μ_i with μ_{it-1}^e in (4) and solving for μ_{it-1}^e yields (6).

$$\mu_{it-1}^e = \beta_{ijt} - X_{it}\tau - \sum_{j=2}^{J} L_{it}^j\delta^j - \varepsilon_{it} \tag{6}$$

Similarly, we can use the lagged version of (6) to derive μ_{it-2}^e. Substituting μ_{it-1}^e and μ_{it-2}^e into (5) and rearranging yields (7).

$$\beta_{ijt} - \beta_{ijt-1} = (X_{it} - X_{it-1})\tau + \sum_{j=2}^{J} L_{it}^j\delta^j - \sum_{j=2}^{J} L_{it-1}^j\delta^j + (\mu_{it-1}^e - \mu_{it-2}^e) + (\varepsilon_{it} - \varepsilon_{it-1}) \tag{7}$$

Note that first differencing does not eliminate worker fixed effects in this case. Therefore, estimating (7) directly does not yield unbiased estimates because the term (μ_{it-1}^e-μ_{it-2}^e) is likely to be correlated with the current job level assignment. As suggested by Gibbons *et al.* (2005), one can use the information from period t-1 and earlier to address this endogeneity problem. With instruments, equation (7) can be estimated using a GMM estimation.

In columns 5 and 6 of Table 5 we report GMM estimation results for equation (7) for our single firm dataset, where column 5 does not include performance ratings while column 6 does. In these tests we do not include a promotion prize variable since instrumenting for the promotion prize is potentially problematic. In column 3 of Table 6 we report GMM estimation results for equation (7) for the Finnish dataset.

In these tests the current job level is treated as an endogenous variable. To instrument for this variable, we use the interaction between job level dummies and salary in t-2, i.e., interactions between L_{it-2}^{j} and w_{it-2}, a binary variable that takes a value of one if the worker was promoted at the end of t-2, and the salary increase from t-2 to t-1.[33] Focusing on Table 5, we see that results are qualitatively unchanged. That is, bonuses rise with job level, rise with job level tenure, and fall with the performance rating. Similarly, in Table 6 we continue to see that bonuses rise with job level and rise with job level tenure at low levels of job level tenure.

Interestingly, the results concerning age are different in these tests. In columns 5 and 6 of Table 5 the coefficients on the age variable are now positive and statistically significant while on the age squared variable the coefficients are negative and statistically significant. In other words, these tests are consistent with the bonus rising with age at young ages. In contrast, in column 3 of Table 6 the coefficient on the age variable and the coefficient on the age squared variable are both statistically insignificant. That is, in this test of the Finnish dataset the results suggest that bonus does not vary with age in a systematic way.[34]

As indicated, the second empirical concern with the fixed effects approach taken in the previous subsection is that it can be problematic if the return to worker ability differs across job levels. Gibbons *et al.* (2005) refer to this as the comparative advantage issue. As discussed above, in our theoretical model equilibrium bonus payments do not directly depend on worker ability. However, if we assumed output was a function of ability times effort rather than ability

[33] To assess the predictive power of our instruments, we perform F-tests for the joint significance of the instruments in predicting the endogenous variables (current job level assignment and also in some later tests the lag of bonus payments). In all cases the instruments are jointly significant at the 1% level.

[34] In the related analyses in Gibbons *et al.* (2005) and Lluis (2005), another variable which these authors refer to as the "skill index" is also included. This variable summarizes the return to observable variables such as education categories, gender, race, and year. For our single firm dataset we also estimated equation (7) including a skill index variable, where the skill index variable was estimated using a regression of bonus payments on job level, age, tenure at current job level, education categories, gender, race, and year controls. There was no effect on the qualitative nature of the results.

plus effort, then bonus payments would depend on worker ability. The following test deals with this issue.

We begin by rewriting equation (4) as follows.

$$\beta_{ijt} = X_{it}\tau + \sum_{j=2}^{J} L_{it}^{j}\delta^{j} + \sum_{j=2}^{J} L_{it}^{j}d^{j}\mu_{i} + \varepsilon_{it} \tag{8}$$

Note that equation (8) incorporates comparative advantage. That is, the return to worker-specific or unobservable ability both vary across job levels, while the return to observables indicated by X_{it} (age, tenure at current job level, and performance ratings) do not depend on job level. It is easy to see that first differencing does not eliminate unobserved worker ability in this case.

Following Gibbons *et al.* (2005), we use a quasi-differenced version of (8) to estimate parameters consistently. Solving for μ_i from (8) yields (9).

$$(1/\sum_{j=2}^{J} L_{it}^{j}d^{j})(\beta_{ijt} - X_{it}\tau - \sum_{j=2}^{J} L_{it}^{j}\delta^{j} - \varepsilon_{it}) \tag{9}$$

A lagged version of (9) now yields (10).

$$\mu_i = (1/\sum_{j=2}^{J} L_{it-1}^{j}d^{j})(\beta_{ijt-1} - X_{it-1}\tau - \sum_{j=2}^{J} L_{it-1}^{j}\delta^{j} - \varepsilon_{it-1}) \tag{10}$$

Let $D_{ijt}=(1/\sum_{j=2}^{J} L_{it}^{j}d^{j})$ and $e_{it}=\varepsilon_{it}-(D_{ijt-1}/D_{ijt})\varepsilon_{it-1}$. Substituting the expression in (10) into (8) now yields (11).

$$\beta_{ijt} - (D_{ijt-1}/D_{ijt})\, \beta_{ijt-1} = X_{it}\tau + \sum_{j=2}^{J} L_{it}^{j}\delta^{j} - (D_{ijt-1}/D_{ijt})(X_{it-1}\tau + \sum_{j=2}^{J} L_{it-1}^{j}\delta^{j}) + e_{it} \tag{11}$$

Because there is no learning, the endogeneity problem arises only from β_{ijt-1}. Using instruments for the lag of bonus payments allows us to consistently estimate the model's parameters.

Using the instruments discussed above, we apply GMM estimation to (11) to estimate the model's parameters. Columns 7 and 8 of Table 5 report GMM estimation results for equation (11) for our single firm dataset, where column 7 does not include performance ratings while column 8 does. As in the GMM tests reported in columns 5 and 6, we do not include a promotion prize variable. In column 4 of Table 6 we report GMM estimation results for equation (11) for the Finnish dataset.[35]

[35] In Appendix B Table B1, we report results for the regression specification used for columns 1 through 4 of Table 5 for the subsamples employed in our learning and comparative advantage tests. In Appendix B Table B2, we report results for the regression specifications used for columns 1 and 2 of Table 6 for the subsamples employed in our learning and comparative advantage tests. The qualitative nature of the results is mostly unchanged.

The results in columns 7 and 8 of Table 5 are qualitatively the same as in columns 5 and 6. That is, bonus payments rise with job level, rise with tenure at level, and fall when the performance rating is higher. Further, these results indicate that the bonus rises with age at young ages. We also find evidence that the return to unobserved ability is higher at levels above one in comparison to the return at level one. But we do not find clear evidence that the return to unobserved ability rises monotonically with the job level.

The results in column 4 of Table 6 are qualitatively similar to the results in the other tests reported in that table. That is, bonus payments rise with job level and rise with tenure at job level at low values for tenure at level. In contrast to column 3, however, and similar to the findings in columns 1 and 2, there is some evidence that bonus payments rise with age at low ages. But the evidence for this finding is much weaker than in columns 1 and 2. Finally, similar to what was true in columns 7 and 8 of Table 5, we find that the return to unobserved ability is higher at levels two and three than at level one, but the evidence that it is higher at level three than at level two is weak.[36]

In the last regression of this subsection we change the manner in which performance ratings enter the specification in our analysis of the Baker, Gibbs, and Holmstrom dataset. In particular, we allow the returns to performance to vary across job levels. In other words, in this specification both unobserved ability and performance are allowed to be rewarded differently across job levels. The basic logic of this regression is that, if the return to unobserved worker ability in terms of higher bonus payments varies with job level which is the maintained assumption in our comparative advantage regressions, then it would seem natural that the return to performance which is likely correlated with unobserved worker ability should also rise with job level.

---

[36] We have also conducted Hansen's J-test for our GMM tests and similar to the findings in Lluis (2005) and Hunnes (2012) which conduct related analyses we find high levels of significance. Gibbons *et al.* (2005) do not report Hansen's J-test for their GMM tests and instead argue that, because of the complicated nonlinear nature of the model, the test should not be thought of as a standard test of the validity of the instruments employed. See Gibbons *et al.* (2002) for a detailed discussion.

To see our approach for this regression more clearly, we rewrite equation (8) as equation (12).

$$\beta_{it} = X_{it}\tau + \sum_{j=2}^{J} L_{it}^{j}\delta^{j} + \sum_{j=2}^{J} L_{it}^{j}\gamma^{j}p_{it} + \sum_{j=2}^{J} L_{it}^{j}d^{j}\mu_{i} + \varepsilon_{it}, \qquad (12)$$

where p_{it} denotes worker i's performance rating in period t. As in the earlier specifications, p_{it} is measured on a three-point scale, where 1 represents the best performance and 3 the worst. Note that to minimize the number of parameters to be estimated, we depart from earlier specifications by not including performance ratings in the form of binary indicator variables.

We employ a quasi-differenced version of (12) which is given in (13).

$$\beta_{ijt} - (D_{ijt-1}/D_{ijt})\beta_{ijt-1} = X_{it}\tau + \sum_{j=2}^{J} L_{it}^{j}\delta^{j} + \sum_{j=2}^{J} L_{it}^{j}\gamma^{j}p_{it} - (D_{ijt-1}/D_{ijt})(X_{it-1}\tau + \sum_{j=2}^{J} L_{it-1}^{j}\delta^{j} + \sum_{j=2}^{J} L_{it-1}^{j}\gamma^{j}p_{it-1}) + e_{it} \qquad (13)$$

Using instruments for the lag of bonus payments, we can consistently estimate the model's parameters.

Results of this estimation are reported in column 9 of Table 5. There are two findings from this regression worth noting. First, the results in terms of the predictions of the theory are qualitatively the same as in the similar regression reported in columns 7 and 8. That is, bonus payments rise with job level, rise with job level tenure, and rise with age at low ages. Second, the coefficients for performance ratings across job levels are significantly different from each other (the p-value for the test of joint equality of these coefficients is 0.0005). Further, viewing these coefficients individually, we see that better performance is associated with larger increases in bonus payments at job levels three and four relative to one and two.

5.4 Trade-Off Between Bonus Payments and Promotion Incentives

The fifth prediction is that there is a trade-off between incentives provided through bonus payments and incentives provided through the possibility of future promotion. Recall that since a worker experiences an increase in expected utility upon being promoted, the probability of being promoted provides a worker with an incentive to exert effort. If the promotion prize is larger which we capture in our theoretical modelling by decreasing discounting, then a smaller bonus rate is required to induce an efficient effort level.[37]

Note that Tables 5 and 6 both include a promotion prize variable and, as the theory predicts, in both tables it is consistently found that a higher promotion prize translates into smaller bonus payments. In this subsection we describe how we construct the promotion prize variable and also provide further tests of this fifth prediction.

To test the prediction that the size of bonus payments is negatively related to the expected promotion prize, we first need to translate the prediction into a specific statement concerning what we observe in the data. The obvious candidate for measuring expected worker utility is expected total compensation, where total compensation refers to the sum of salary plus any bonus payment. However, for a worker who is not promoted in a given year we do not observe what compensation would have been if a promotion had taken place.

Let Δ_{it}^{e} be worker i's expected promotion wage increase in period t, where Δ_{it}^{e} is defined by equations (14) and (15). Note, in (14) and (15) below $prom_{it}=1$ means worker i is promoted in period t and $prom_{it}=0$ means the worker is not promoted.

$$\Delta_{it}^{e} = C_{it+1}^{P} - C_{it} \text{ if } prom_{it}=1 \tag{14}$$

$$\Delta_{it}^{e} = C_{it+1}^{P,e} - C_{it} \text{ if } prom_{it}=0 \tag{15}$$

In equations (14) and (15) C_{it+1}^{P} and C_{it} denote worker i's compensation in period t+1 when the worker is promoted at the end of period t and worker i's compensation in period t, respectively. For workers not promoted C_{it+1}^{P} is not observed and must be predicted. The variable $C_{it+1}^{P,e}$ is the predicted value for C_{it+1}^{P} for workers not promoted.

[37] Corollary 5 considers what happens in our theoretical model when there is a change that affects the size of the promotion prize but no change in how effort affects productivity in the low level job. The result is no change in equilibrium effort which the model predicts is always at the efficient level and a decrease in the size of bonus payments. Note that we are unable to test for the prediction concerning equilibrium effort since we do not have measures of effort levels. Also, in testing the fifth prediction we control for most of the factors that affect efficient effort choice in the theoretical model, i.e., job level, job level tenure, and worker experience which we proxy with worker age. It would be optimal to also control for firm tenure which would matter in a richer specification in which turnover arises in equilibrium. But our measure for firm tenure in the single firm dataset is noisy (see the related discussion in footnote 18) and so we do not include this variable.

We employ an approach similar to one employed in DeVaro and Waldman (2012) in analysing a related problem to construct expected promotion wage increases. In particular, we take into account worker heterogeneity by employing a detailed set of control variables in constructing expected promotion prizes. In the first step we estimate equation (16) for the subsample of observations in which promotion occurred, where Y_{it}^{P} is a vector of control variables.

$$C_{it+1}^{P} - C_{it} = Y_{it}^{P}\kappa_{Y} + \psi_{it} \tag{16}$$

For each non-promotion observation, we then construct an expected promotion wage increase by employing the values for the control variables for the observation and the estimated coefficients from our estimation of equation (16).

We use three different sets of control variables in estimating equation (16). The first set of control variables consists of job level, tenure at current level, year, and worker fixed effects. Using these control variables we construct what we call Promotion Wage Increase A for each observation. For the second set of control variables we add to the prior list gender, race, age, tenure at the firm, and education level (worker fixed effects are removed) and we call the result Promotion Wage Increase B. Finally, for our third set of control variables we add to the second set just described job titles. The resulting predicted promotion prizes are called Promotion Wage Increase C.[38,39]

---

[38] The two standard approaches for the estimation of endogenous treatment effects are instrumental variables and control function procedures (see Robinson (1989) for a discussion and references concerning this issue in the union membership context). To apply these techniques to our problem, however, would require a variable that is correlated with the probability of earning a promotion but is uncorrelated with the size of compensation changes. Since the probability of earning a promotion and the size of compensation changes are both determined to a great extent by worker performance, finding a variable with the required properties is a difficult task. One variable that would satisfy the required conditions is the separation decision of a worker in a higher managerial position. That is, such a separation can increase the probability of promotion for lower level workers who can potentially fill the now open position, but is likely uncorrelated with compensation increases for any promoted worker and also those not promoted. Unfortunately, this information is not available in our datasets.

[39] To address the endogeneity problem concerning who gets promoted it would be useful to include performance ratings in our construction of promotion wage increases and also expected promotion prizes. Because the Finnish dataset does not include performance ratings, we do not include performance ratings in the construction of the promotion wage increase variables or the promotion prize variables. However, we have conducted tests on the Baker, Gibbs, and Holmstrom dataset where we include performance ratings in the construction of promotion wage increase variables and promotion prize variables and there was no change in the qualitative nature of the results.

To incorporate the effect of promotions, let I_{it}^* denote a latent index variable such that worker i is promoted, i.e., $prom_{it}=1$, if and only if $I_{it}^* \geq 0$. We estimate a logit regression of the following form.

$$I_{it}^* = \tau_1 X_{it} + \upsilon_{it}, \tag{17}$$

where X_{it} is a vector of control variables and it includes the worker's gender, job level, tenure at the current job level, age, and indicator variables for education categories. Using parameter estimates from (17), we then derive an expected promotion probability for each observation, $prom_{it}^e$.

Finally, using the expected promotion probability for each observation and the three estimates for promotion wage increases, we construct three estimates for expected promotion prizes which we denote Promotion Prize A, Promotion Prize B, and Promotion Prize C. Specifically, this construction is given in equation (18).

$$\Psi_{it}^e = prom_{it}^e \times \Delta_{it}^e \tag{18}$$

In Tables 5 and 6 we estimated equation (2) and employed Promotion Prize A as our promotion prize variable. We found, consistent with the theory, that higher promotion prizes are associated with smaller bonus payments. In Tables 7 and 8 we again estimate equation (2) where our focus now is whether this conclusion is robust to employing alternative measures of the promotion prize and varying the number of controls. Note that since equation (2) includes a predicted variable, Ψ_{it}^e, as an independent variable, conventional methods underestimate standard errors (see Murphy and Topel (1985) for a discussion). Therefore, in Tables 7 and 8 we adjust standard errors to take into account the sampling variability of this term. In particular, we implement a non-parametric bootstrap method which allows us to use the variation in the bootstrapped estimates of ρ to adjust the standard error estimated from the original sample.[40]

---

[40] The method we implement is very similar to the approach that is used to compute standard errors with multiple imputed data (see Rubin (1987)). It can be summarized as follows. Drawing independent random samples from the subsamples of promoted and non-promoted workers, respectively, we first generate 50 datasets in addition to the original one. Then, we estimate (2) for each bootstrap sample and save the results. The corrected standard error is given by the formula $(s_P^2+\sigma_P^2)^{1/2}$, where s_P^2 is the sample variance estimated from the original sample and σ_P^2 is the variance of the point estimates across the bootstrap samples. Note that correcting the standard errors using the method just described does not affect the qualitative nature of the results.

Table 7 reports results for the single firm dataset. The top panel of Table 7 reports results when we employ Promotion Prize A, while the middle panel reports results for Promotion Prize B and the bottom panel shows results for Promotion Prize C. In the first column of each panel we include no controls that are in addition to the Promotion Prize variable except for the standard controls that consist of age, race, gender, education level, job level, and year. In the second column we add tenure at the current job level while in the third column we add average salary increase at the current job level but do not control for tenure at level. In the fourth column we include both average salary increase at the current job level and job level tenure. In the fifth column the only added variable is performance ratings, while in the sixth column we add tenure at current level, average salary increase at current level, and performance ratings. We consistently find that, as predicted by our theoretical model, the coefficient on the promotion prize variable is negative and statistically significant at the 1% level.[41]

To get a sense of the magnitude of the effect, consider the top panel which reports results for Promotion Prize A (the results for the other promotion prizes reported in the lower panels are similar). In the first column which does not include any additional controls, the coefficient on the promotion prize variable is -0.180 which means that a one dollar increase in the expected promotion prize leads to a 0.18 dollar decrease in bonus payments. In columns 2 through 6 we include additional controls and the result is that the coefficient rises in absolute value, where the largest effects are in columns 5 and 6 that include performance ratings as control variables.

In Table 8 we conduct the same tests using the multi-firm dataset and the results are similar (note that this table does not include tests like in columns 5 and 6 of Table 7 because the multi-firm dataset does not include performance ratings). The only noticeable difference is that the absolute values for the coefficients are much larger in Table 8 than in 7.[42]

[41] The regressions reported in Tables 7 and 8 do not include worker fixed effects. We continue to find results consistent with the theoretical prediction, however, when worker fixed effects are included.

[42] Related to footnote 37, the theoretical prediction can be interpreted as bonus size should be negatively related to the size of the promotion prize, holding effort fixed. One could argue, therefore, that in testing our fifth prediction there should be controls for worker effort. From this perspective columns 5 and 6 of Table 7 could be considered superior tests of the theory.

We believe the explanation for the difference in the absolute values of the coefficients across Tables 7 and 8 is related to a discussion above concerning why the first four columns of Tables 5 and the first two columns of 6 are different in terms of the effect of age on bonus payments. In the Baker, Gibbs, and Holmstrom firm stock holding and stock options are likely important. When the expected promotion prize rises, our theory predicts that bonus payments and stock based incentives should fall to keep effort at the first best level. We only observe part of this effect in the tests on the Baker, Gibbs, and Holmstrom dataset because we only have data on bonuses. For the Finnish firms in our multi-firm dataset stock based compensation is rare. So the coefficients in Table 8 should be larger in absolute value because those coefficients should reflect close to the full decrease in bonus and stock based compensation when the promotion prize rises.

6. Conclusion

One way in which firms frequently provide workers with an incentive for effort is through the use of bonus contracts. In this paper we have focused both theoretically and empirically on understanding the determinants of the size of bonuses and, in particular, our focus has been on bonuses in a setting characterized by promotion tournaments as first analysed by Lazear and Rosen in their seminal 1981 paper. In previous literature focused on promotion tournaments no distinction is typically made between salary based compensation and bonus based compensation. We extend the tournament literature to capture this distinction and then empirically investigate the resulting testable implications.

In constructing a tournament model which makes a distinction between salary and bonus payments we employ a hybrid approach that combines elements of the classic tournament approach found in Lazear and Rosen's (1981) paper and the market-based approach first explored in Gibbs (1995) and Zabojnik and Bernhardt (2001). In our model, at the beginning of each period the firm commits to a compensation contract consisting of a salary, minimum output level required to achieve a bonus, and a bonus rate. This is like the classic approach in the sense that the firm has some commitment ability in determining compensation. But the compensation increase that follows a promotion is due to the signalling role of promotion which is the approach taken in the market-based approach.

Our theoretical analysis yields five predictions. First, bonus size should increase with job level holding job level tenure and worker age fixed, where the logic of this prediction is that the return to worker effort increases with job level in our model so the efficient effort level increases with job level. Second, bonus size should increase with job level tenure holding job level and age fixed. This follows given our assumption of task specific human capital. Third, holding job level and job level tenure fixed, bonus size should increase with worker age. One reason is that higher age means more human capital accumulation and this increases the efficient effort level. Fourth, bonus size should increase with performance holding fixed job level, job level tenure, and age. Fifth, bonus size is negatively related to the size of expected promotion prizes. Here the argument is that in aggregate bonus incentives plus promotion incentives in equilibrium achieve efficient effort levels. So if expected promotion prizes are larger, then smaller bonuses are needed to achieve efficient levels.

After developing these five predictions, we provide an empirical analysis using two distinct datasets. The first is the dataset first employed in Baker *et al*. (1994a,b) classic empirical study of wage and promotion dynamics in the financial services industry. The second is a multi-firm dataset that covers most white collar manufacturing employment in Finland during the time period 2003 to 2012. Our empirical analysis provides strong support for four of the five predictions and mixed support for the fifth. The prediction with mixed support is the one concerning bonus size being positively related to age. Focusing on the results in our fixed effects specifications, this prediction is not supported in the Baker, Gibbs, and Holmstrom dataset while in the Finnish dataset it is supported for workers up to their mid 40s to mid 50s. As we discussed in detail earlier, we feel this finding in the Finnish dataset is in fact consistent with our theoretical approach given existing empirical evidence concerning Finland which suggests that human capital typically peaks in that labour force when workers are in their early to mid 50s.

As we also discussed earlier, the lack of support for the age prediction in the Baker, Gibbs, and Holmstrom dataset may be due to the importance of employee stock holdings and stock options in that firm.

There are a number of directions in which the analysis presented here could be extended. For example, we think it would be interesting to formally extend both the theoretical and empirical analyses by incorporating stock ownership and stock options. Many firms provide incentives at higher job levels through stock ownership and stock options. Related to our discussion concerning the results in the Baker, Gibbs, and Holmstrom dataset concerning the age prediction, there should thus be a trade-off between incentives provided through stock ownership and stock options and bonus size similar to the trade-off focused on in this paper between bonus size and promotion incentives. We think it would be of interest to add this third avenue through which incentives can be provided into a promotion tournament type setting and formally investigate both theoretically and empirically the predictions that result.

Another direction of interest would be to extend the theory to identify how different firm attributes affect the use of bonuses in our promotion tournament framework. We could then investigate the resulting theoretical predictions using the multi-firm Finnish dataset. And we also made a number of simplifying assumptions to keep the theoretical model tractable and it might be fruitful to relax some of these assumptions to the extent possible. This includes formally extending the analysis to more periods and allowing for worker risk aversion.

Appendix A: Proofs

In the Appendix we provide proofs of the lemmas, propositions, and corollaries in Section 2 and 3. In the proof of Proposition 1 we also provide the parameter restrictions that guarantee the conditions described at the beginning of Subsection 2.2. Note also that due to space considerations proofs are somewhat abbreviated.

Proof of Lemma 1: In the proof of Proposition 1 we show that for each group k, k=1,2, there is a cutoff ability level θ_k^+, $\theta_k^+>\theta'$, such that worker i in group k is assigned to job k+1 in period 2 by the first period employer if $\theta_i\geq\theta_k^+$ and is assigned to job k if $\theta_i<\theta_k^+$. For the proof of Lemma 1 we take this result as given.

Consider group k and the compensation determination process at the beginning of period 2. Given our trembling hand type assumption, the market contract offer will be consistent with zero profits under the assumption that a worker who moves is the lowest ability type among workers with the same labour market signal or job assignment, i.e., θ_k^L for workers not promoted and θ_k^+ for promoted workers. The logic is that this is the worker for whom the initial employer's foregone profits from making the mistake of not matching is the lowest. Further, this worker would be assigned to the same job the initial employer assigned the worker to if the worker were to move given $\theta_k^+>\theta_k'$ (this follows given a parameter restriction as discussed in the proof of Proposition 1). And also, competition among firms means the market contract offer will be the one that maximizes the utility of such a worker given the zero expected profit constraint. Given everyone is risk neutral, this yields that for workers initially assigned to job k the market offers a salary equal to $h_2(c_k+d_k\theta_k^L)$-z, the minimum output specified in the contract is the same value, and the bonus rate equals one, while for workers initially assigned to job k+1 the market offers a salary equal to $h_1(c_{k+1}+d_{k+1}\theta_k^+)$-z, the minimum output specified in the contract is the same value, and the bonus rate equals one.

Given the presence of firm specific human capital, i.e., $s_1>h_1$ and $s_2>h_2$, the initial employer always matches which means the utility of any worker assigned to job j from staying just equals the utility associated with the worker leaving. This is the second condition in iii) of the lemma given that efforts are chosen efficiently. The first condition in iii) follows from our assumption that the lower bound on the bonus equals zero (see footnote 16).

Finally, the initial employer will want to maximize second period profits in choosing the compensation contract for each worker given that the worker's utility associated with the contract just matches the worker's utility from accepting the market wage offer. As in any standard agency problem with risk neutrality this means the bonus rate is set equal to one and the worker chooses the efficient effort level. This proves i), ii), and iii).

Proof of Lemma 2: Given our focus is pure strategy Perfect Bayesian equilibria and given there is no stochastic element in the production functions, at the end of period 1 upon observing a worker's first period output a worker's first period employer and also the worker learn the worker's innate ability level with certainty. Also, at the beginning of period 2 other firms know this. It can also be shown that any specific belief about a worker's innate ability at the end of period 1 translates into unique behaviour in period 2 which in combination with the previous results means that the specific contracts signed in period 1 have no effect on period 2 behaviour.

We know from iii) of Lemma 1 and that there is firm specific human capital that hiring a worker in period 1 is associated with strictly positive expected profits in period 2. Competition among employers in hiring in period 1 thus yields that for workers in each group k expected compensation must exceed expected first period output. This proves ii) given a single contract is offered to workers in each group k in period 1 (and given $e_{ijk1}=e_j^*(1)$ which is proven below).

If $\theta_i \geq \theta_k'$, then given the market contracts derived in the proof of Lemma 1 it must be the case that $U^{2M}(\theta_{ik+1k2})>U^{2M}(\theta_{ikk2})$ (this is shown formally in the proof of Proposition 1). Given $\theta_k^+>\theta_k'$ for all k, k=1,2, we now have that $U^{2M}(\theta_{ik+1k2})>U^{2M}(\theta_{ikk2})$ if $\theta_i=\theta_k^+$.

Now consider the first period choice of effort of worker i in group k. This worker chooses first period effort to maximize the worker's expected discounted utility over the two periods which yields the first order condition $\beta_{kk1}d_k+\delta(\partial EU_{i2}(e_{i1})/\partial e_{i1})=g'(e_{i1})$, where $EU_{i2}(e_{i1})$ is the worker's expected utility in period 2 as a function of the first period effort choice. Let e_{1k}^* be the equilibrium effort choice in period 1 for workers in group k. In equilibrium this condition reduces to $\beta_{kk1}d_k+\delta(\partial EU_{i2}(e_{1k}^*)/\partial e_{i1})=g'(e_{1k}^*)$. We know that increasing first period effort increases the first period employer's belief concerning the worker's innate ability. Given the market contracts and iii) of Lemma 1, this only changes second period utility if the initial job assignment changes, i.e., second period utility is only affected if the increase in effort causes the initial employer's belief concerning the worker's ability to go from below θ_k^+ to be equal to or above θ_k^+. Further, since in equilibrium this belief is correct, $\theta_k^+>\theta_k'$, and the market contract offers, we have that second period utility rises when an increase in effort causes the job assignment to change. This means $\partial EU_{i2}(e_{1k}*)/\partial e_i>0$ which, in turn, yields $\beta_{kk1}d_k<g'(e_{1k}*)$.

By definition, we know $d_k=g'(e_k*(1))$. Given this, now consider the equilibrium contract in period 1 for group k and the resulting choice of a first period effort level. The contract must maximize a group k worker's expected utility over the two periods subject to a zero expected profit constraint. But we know that the choice of a contract has no effect on second period expected utility so the contract must maximize first period expected utility subject to a zero expected profit constraint. Suppose $e_{1k}*\neq e_k*(1)$. Then there would be an alternative contract that results in the worker choosing $e_k*(1)$, that satisfies zero expected profits, and that achieves higher expected worker utility which contradicts $e_{1k}*\neq e_k*(1)$. So $e_{1k}*=e_k*(1)$. But given $\beta_{kk1}d_k<g'(e_{1k}*)$ and $d_k=g'(e_k*(1))$ we have $\beta_{kk1}<1$ for all k, k=1,2. This proves i).

Proof of Proposition 1: Later in the proof we provide the parameter restrictions that guarantee that group 1 workers are assigned to job 1 in period 1 and in period 2 are assigned to either job 1 or job 2, group 2 workers are assigned to job 2 in period 1 and either job 2 or job 3 in period 2, and there exist values θ_1^+ and θ_2^+, $\theta_1^+>\theta_1'$ and $\theta_2^+>\theta_2'$, such that in period 2 a worker in group k, k=1,2, with innate ability $\theta_i\geq\theta_k^+$ is assigned to job k+1 while a worker with innate ability $\theta_i<\theta_k^+$ is assigned to job k. Taking these conditions as given, iii) follows from the arguments in the proof of Lemma 1.

The next step is to provide the parameter restrictions that guarantee that group k workers, k=1,2, are assigned to job k in period 1. As stated earlier, given our focus is pure strategy Perfect Bayesian equilibria and given there is no stochastic element in the production functions, at the end of period 1 upon observing a worker's first period output a worker's first period employer learns the worker's innate ability level with certainty. Also, at the beginning of period 2 other firms know this and it can also be shown that any specific belief about a worker's innate ability results in a unique set of behaviours in period 2. Combining this result with the idea that competition for workers in period 1 means that period 1 equilibrium contracts and job assignments maximize expected worker utility subject to a zero expected profit constraint yields that the period 1 job assignment must maximize expected period 1 surplus. So workers in group 1 are assigned to job 1 in period 1 as long as $c_1+d_1(E(\theta^1)+e_1{}^*(1))-g(e_1{}^*(1))>\max\{c_2+d_2(E(\theta^1)+e_2{}^*(1))-g(e_2{}^*(1)),c_3+d_3(E(\theta^1)+e_3{}^*(1))-g(e_3{}^*(1))$. Similarly, workers in group 2 are assigned to job 2 in period 1 as long as $c_2+d_2(E(\theta^2)+e_2{}^*(1))-g(e_2{}^*(1))>\max\{c_1+d_1(E(\theta^2)+e_1{}^*(1))-g(e_1{}^*(1)),c_3+d_3(E(\theta^2)+e_3{}^*(1))-g(e_3{}^*(1))$. In turn, with these parameter restrictions i) follows from arguments in the proof of Lemma 2.

We now compare efficient assignment rules in period 2 for the period 1 employer and for prospective employers. As indicated earlier, for the first period employer the efficient assignment rule for period 2 for a worker in group k is to promote the worker when $\theta_i\geq\theta_k'$ and not promote the worker when $\theta_i<\theta_k'$, where θ_k' satisfies $s_2[c_k+d_k(\theta_k'+e_k{}^*(s_2))]-g(e_k{}^*(s_2))=s_1[c_{k+1}+d_{k+1}(\theta_k'+e_{k+1}{}^*(s_1))]-g(e_{k+1}{}^*(s_1))-z$. This follows given our assumption that $s_1d_{k+1}>s_2d_k$ for all k, k=1,2. For a prospective employer the efficient assignment rule is to assign the worker to job k+1 when $\theta_i\geq\theta_k{}^{M\prime}$ and to job k when $\theta_i<\theta_k{}^{M\prime}$, where $\theta_k{}^{M\prime}$ satisfies $h_2[c_k+d_k(\theta_k{}^{M\prime}+e_k{}^*(h_2))]-g(e_k{}^*(h_2))=h_1[c_{k+1}+d_{k+1}(\theta_k{}^{M\prime}+e_{k+1}{}^*(h_1))]-g(e_{k+1}{}^*(h_1))$. This follows given our assumption that $h_1d_{k+1}>h_2d_k$. Note that θ_k' increases with z while $\theta_k{}^{M\prime}$ does not and we assume that z is sufficiently large that $\theta_k'>\theta_k{}^{M\prime}$ which is a sufficient condition for the model to exhibit a promotion signaling distortion, i.e., the proportions of workers promoted by the first period employer are below the efficient levels.[43]

The next step is to show that there exist values $\theta_1{}^+$ and $\theta_2{}^+$ such that $\theta_1{}^+>\theta_1'$ and $\theta_2{}^+>\theta_2'$ and ii) holds. In this step of the proof we assume no demotions and no promotions of more than one level. Call $\theta_{jk}{}^L$ the lowest ability worker in group k assigned to job j by the first period employer in period 2. Consider for the moment group 1 in which case our focus is $\theta_{11}{}^L$ and $\theta_{21}{}^L$.

[43] If $\theta_k'\leq\theta_k{}^{M\prime}$, then promotions serve as signals but there is no promotion signaling distortion for reasons similar to those found in Golan (2005) in a related analysis. A related discussion appears in the text following Proposition 1.

Suppose they are both below $\theta_1^{M\prime}$. As shown below, in this case the market contract offers are such that a worker who moves receives the same utility whether the worker was assigned to job 1 or job 2 by the first period employer, so the market utility that must be matched is independent of the initial job assignment. But this means workers will be assigned efficiently in which case $\theta_1^{+}=\theta_1'$ which contradicts the supposition.

Suppose $\theta_{11}^{L}<\theta_1^{M\prime}$ and $\theta_{21}^{L}<\theta_1^{M\prime}$. The logic for why market contract offers are such that a group 1 worker who moves receives the same utility whether the worker was assigned to job 1 or job 2 by the first period employer is as follows. First, consider a worker in group 1 with ability $\theta^{\#}$ assigned to job 1 by the first period employer in period 2. Given $\theta_{11}^{L}<\theta_1^{M\prime}$, arguments like those in the proof of Lemma 1 yield that the market offers to assign the worker to job 1, offers a salary equal to $h_2(c_1+d_1\theta_{11}^{L})$-z, the minimum output specified in the contract is the same value, and the bonus rate equals one. If the worker moves the worker chooses the efficient effort level, $e_1^*(h_2)$, given the bonus rate equals one. This means the worker's utility associated with moving equals $h_2(c_1+d_1\theta_{11}^{L})-z+h_2(d_1(\theta^{\#}-\theta_{11}^{L})+e_1^*(h_2))-g(e_1^*(h_2))=h_2(c_1+d_1\theta^{\#}+e_1^*(h_2))-z-g(e_1^*(h_2))$.

Now consider a worker in group 1 with ability $\theta^{\#}$ assigned to job 2 by the first period employer in period 2. Given $\theta_{21}^{L}<\theta_1^{M\prime}$, arguments like those in the proof of Lemma 1 yield that the market offers to assign the worker to job 1, offers a salary equal to $h_2(c_1+d_1\theta_{21}^{L})$-z, the minimum output specified in the contract is the same value, and the bonus rate equals one. If the worker moves, we again have the worker chooses $e_1^*(h_2)$. This worker's utility associated with moving thus equals $h_2(c_1+d_1\theta_{21}^{L})-z+h_2(d_1(\theta^{\#}-\theta_{21}^{L})+e_1^*(h_2))-g(e_1^*(h_2))=h_2(c_1+d_1\theta^{\#}+e_1^*(h_2))-z-g(e_1^*(h_2))$. So a worker who moves receives the same utility whether the worker was assigned to job 1 or job 2 by the first period employer.

Suppose $\theta_{11}^{L}\geq\theta_1^{M\prime}$ and $\theta_{21}^{L}<\theta_1^{M\prime}$. Then the market contract offered to workers assigned to job 1 by the first period employer is consistent with the worker being assigned to job 2 by an alternative employer while the market contract offered to workers assigned to job 2 by the first period employer is consistent with the worker being assigned to job 1 by an alternative employer.

We also know $\theta_{21}^L=\theta_1^L$ and $\theta_{21}^L<\theta_{11}^L$ in this case. But given our assumptions $s_1d_2-s_2d_1>h_1d_2-h_2d_1$ (see also footnote 15), if the firm finds it profitable to assign a θ_1^L worker to job 2 then it must also find it profitable to assign a θ_{11}^L worker to job 2 which is a contradiction.

The only other possibility is that $\theta_{11}^L<\theta_1^{M\prime}$ and $\theta_{21}^L\geq\theta_1^{M\prime}$. Suppose $\theta_{21}^L=\theta_1^{M\prime}$. Then the market contract offers are such that if a $\theta_1^{M\prime}$ worker moves the worker receives the same utility whether the worker was assigned to job 1 or job 2 by the first period employer. But this means the first period employer should assign the worker to job 1 which is a contradiction. So we have $\theta_{11}^L<\theta_1^{M\prime}$ and $\theta_{21}^L>\theta_1^{M\prime}$. But this means that if a θ_{21}^L worker moves the worker receives higher utility when the worker was assigned to job 2 rather than job 1 by the first period employer. And this, in turn, means that $\theta_{21}^L>\theta_1'$ since otherwise the firm would have an incentive to assign the worker to job 1. Finally, given our assumptions $s_1d_2-s_2d_1>h_1d_2-h_2d_1$ (see also footnote 15), if the firm has an incentive to assign a θ_{21}^L worker to job 2 then it also has an incentive to assign any group 1 worker with higher innate ability to job 2. Thus, there exists a value θ_1^+ with the specified properties. A similar argument yields that there is also a value θ_2^+ with the specified properties.

The last step of the proof is to provide the parameter restrictions such that there are no demotions and no promotions in which the assignment increases by two levels. Clearly there are no demotions for group 1 workers and no promotions in which the assignment increases by two levels for group 2 workers. Consider first group 1 and the idea that there are no two-level promotions. Using logic like that above that showed that $\theta_1^+>\theta_1'$ can be used to show that there is a critical value for promotion to job 3 which is above θ_2'. But by assumption $\theta_1^H<\theta_2'$ so there are no two-level promotions, i.e., no additional parameter restrictions are required to rule out two-level promotions.

Now consider group 2 and the idea that there are no demotions. To rule out demotions we assume parameters are such that $\theta_1^{M\prime}<\theta_2^L$. With this restriction if a group 2 worker is assigned to job 1 in period 2 rather than job 2, the market contract offer is unchanged so there is no effect on the utility of the worker if he or she moves. This means the utility the first period employer needs to match is independent of whether the assignment is to job 1 or to job 2. So in deciding whether to assign the worker to job 1 or job 2 at the beginning of period 2, given $\theta_2^L>\theta_1'$, the firm would prefer to assign the worker to job 2 rather than job 1 which means there are no demotions in equilibrium. This completes the proof.

Proof of Corollary 1: From Lemma 1 we have that, if j=k, then the bonus size for old workers equals $s_2d_je_j*(s_2)$. Since $d_2>d_1$ and $e_2*(s_2)>e_1*(s_1)$, we have that old workers on job level 2 with one period of prior experience on the job level have higher bonuses than old workers on job level 1 with one period of prior experience on the job level. From Lemma 1 we also have that, if j=k+1, then the bonus size for old workers equals $s_1d_j(e_j*(s_1))$. Since $d_3>d_2$ and $e_3*(s_1)>e_2*(s_1)$, we have that old workers on job level 3 with zero periods of prior experience on the job level have higher bonuses than old workers on job level 2 with zero periods of prior experience on the job level. This proves the first part of Lemma 1.

Now consider period 1 and group k. From the proof of Proposition 1 we know that $\beta_{kk1}d_k+\delta(\partial EU_{k2}/\partial e_{k1})=c'(e_k*(1))=d_k$. As δ approaches zero this yields that β_{kk1} approaches one. This, in turn, yields that the expected bonus payment for group k workers in period 1 on job level k approaches $d_k[E(\theta^k)-\theta_k^L+e_k*(1)]$. Since $d_2>d_1$, $e_2*(1)>e_1*(1)$, and $E(\theta^2)-\theta_2^L\geq E(\theta^1)-\theta_1^L$, we know $d_2[E(\theta^2)-\theta_2^L+e_2*(1)]>d_1[E(\theta^1)-\theta_1^L+e_1*(1)]$. This proves the second part of Corollary 1.

Proof of Corollary 2: From Lemma 1 we have that old workers on job level 2 with zero periods of firm level tenure earn a bonus equal to $s_1d_2e_2*(s_1)$, while old workers on job level 2 with one period of firm level tenure earn a bonus equal to $s_2d_2e_2*(s_2)$. Since $s_2>s_1$ and $e_2*(s_2)>e_2*(s_1)$, we have that for old workers on job level 2 the average bonus payment increases with job level tenure.

Proof of Corollary 3: For workers on job level 2, holding job level tenure constant, the only variation involving worker age concerns group 2 workers assigned to job 2 in period 1 and group

1 workers assigned to job 2 in period 2 each of whom has zero prior periods on the job level. From Lemma 1 we have that the bonus for old workers on job level 2 for whom this is the first period on the level equals $s_1 d_2 e_2^*(s_1)$. From Lemma 2 we have that the expected bonus for young workers on level 2 for whom this is the first period on the level is less than $d_2[E(\theta^2)-\theta_2^L+e_2^*(1)]$. We know $s_1>1$ and $e_2^*(s_1)>e_2^*(1)$, so a comparison of these expressions tells us that, if s_1 is sufficiently large, then for workers on job level 2 the average bonus payment increases with worker age given job level tenure is held constant.

Proof of Corollary 4: Consider period 1 and worker i in group k. From the proof of Lemma 2 we know that $\beta_{kk1}d_k+\delta(\partial EU_{i2}/\partial e_{i1})=g'(e_k^*(1))=d_k$. Given δ sufficiently small (as is indicated in the set-up of the model), this equation yields $\beta_{kk1}>0$. Given this logic holds for each k, k=1,2, we now have that the bonus rate in period 1 is positive for each job level. But a positive bonus rate for each job level in period 1 immediately yields that for young workers bonus payments increase with output once job level is held fixed (job level tenure does not vary among young workers).

Proof of Corollary 5: Consider period 1 and worker i in group k. From the proof of Lemma 2 we know that $\beta_{kk1}d_k+\delta(\partial EU_{i2}/\partial e_{i1})=g'(e_k^*(1))=d_k$. We also know from the proof of Lemma 2 that a change in δ does not change equilibrium behaviour in period 2. So an increase in δ increases $\delta(\partial EU_{i2}/\partial e_{i1})$ which, given the equation above means β_{kk1} decreases. Since $e_k^*(1)$ is unchanged with an increase in δ and all group k workers are assigned to job level k in period 1 independent of δ, we now have that the average bonus payment in period 1 decreases for young workers on job level j, j=1,2.

Appendix B: Additional Tables

Universidad Carlos III de Madrid
The Research Institute of the Finnish Economy
Johnson Graduate School of Management, Cornell University

REFERENCES

Asplund, R. (2001), "Finland," in C. Harmon, I. Walker, and N. Westergaerd-Nielsen (eds.), *Education and Earnings in Europe: A Cross-Country Analysis of the Returns to Education*, Edward Elgar Publishing, pp. 68-83.

Baker, G., M. Gibbs, and B. Holmstrom (1994a), "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 109, pp. 881-919.

Baker, G., M. Gibbs, and B. Holmstrom (1994b), "The Wage Policy at a Firm," *Quarterly Journal of Economics*, 109, pp. 921-955.

Baker, G., M.C. Jensen, and K.J. Murphy (1988), "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, 15, pp. 338-375.

Barron, J.M., M.C. Berger, and D.A. Black (2006), "Selective Counteroffers," *Journal of Labor Economics*, 24, pp. 385-409.

Belzil, C. and M.L. Bognanno (2010), "The Promotion Dynamics of American Executives," *Research in Labor Economics*, 30, pp. 189-231.

Benabou, R. and J. Tirole (2016), "Bonus Culture, Competitive Pay, Screening and Multitasking," *Journal of Political Economy*, 124, pp. 305-370.

Bernhardt, D. (1995), "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, pp. 315-339.

Bognanno, M. and E. Melero (2016), "Promotion Signals, Experience, and Education," *Journal of Economics and Management Strategy*, 25, pp. 111-132.

Boschmans, K. (2008), Selected Essays on Personnel Economics: Turnover, Promotions and Incentives," Unpublished PhD Dissertation, University of Antwerp.

Brown, J.N. (1989), "Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed Within Firms," *American Economic Review*, 79, pp. 971-991.

Cassidy, H., J. DeVaro, and A. Kauhanen (2016), "Promotion Signaling, Gender, and Turnover: New Theory and Evidence," *Journal of Economic Behavior and Organization*, 126 (Part A), pp. 140-166.

Cook, J.B. and R.K. Mansfield (2016), "Task-Specific Experience and Task-Specific Talent: Decomposing the Productivity of High School Teachers," *Journal of Public Economics*, 140, pp. 51-72.

Dato, S., A. Grunewald, M. Krakel, and D. Muller (2016), "Asymmetric Employer Information, Promotions, and the Wage Policy of Firms," *Games and Economic Behavior*, 100, pp. 273-300.

DeAngelo, G. and E. Owens (2017), "Learning the Ropes: Task Specific Experience and the Output of Idaho State Troopers," *Journal of Economic Behavior and Organization*, 142, pp. 368-377.

DeVaro, J., S. Ghosh, and C. Zoghi (2018), "Job Characteristics and Labor Market Discrimination in Promotions," *Industrial Relations,* 57, pp. 389-434.

DeVaro, J. and A. Kauhanen (2016), "An 'Opposing Responses' Test of Classic versus Market-Based Promotion Tournaments," *Journal of Labor Economics*, 34, pp. 747-779.

DeVaro, J. and M. Waldman (2012), "The Signaling Role of Promotions: Further Theory and Empirical Evidence," *Journal of Labor Economics*, 31, pp. 227-269.

Fairburn, J.A. and J.M. Malcomson (2001), "Performance, Promotion, and the Peter Principle," *Review of Economic Studies*, 68, pp. 45-66.

Fama, E.F. (1980), "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88, pp. 288-307.

Gathmann, C. and U. Schönberg (2010), "How General is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 28, pp. 1-49.

Ghosh, S. and M. Waldman (2010), "Standard Promotion Practices Versus Up-or-Out Contracts," *Rand Journal of Economics*, 41, pp. 301-325.

Gibbons, R. and K.J. Murphy (1992), "Optimal Incentive Contracts in the Presence of Career Concerns – Theory and Evidence," *Journal of Political Economy*, 100, pp. 468-505.

Gibbons, R. and L. Katz (1991), "Layoffs and Lemons," *Journal of Labor Economics*, 9, pp. 351-380.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2002), "Comparative Advantage, Learning, and Sectoral Wage Determination," NBER Working Paper, No. 8889.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2005), "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 23, pp. 681-724.

Gibbons, R. and M. Waldman (2004), "Task-Specific Human Capital," *American Economic Review*, 94, pp. 203-207.

Gibbons, R. and M. Waldman (2006), "Enriching a Theory of Wage and Promotion Dynamics Inside Firms," *Journal of Labor Economics*, 24, pp. 59-107.

Gibbs, M. (1995), "Incentive Compensation in a Corporate Hierarchy," *Journal of Accounting and Economics*, 19, pp. 247-277.

Gibbs, M. (2003), "Evidence on Wage and Promotion Dynamics in an Internal Labor Market," Mimeo, University of Chicago.

Golan, L. (2005), "Counteroffers and Efficiency in Labor Markets with Asymmetric Information," *Journal of Labor Economics*, 23, pp. 373-393.

Green, J.R. and N. Stokey (1983), "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 91, pp. 349-364.

Greenwald, B. (1986), "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53, pp. 325-347.

Gürtler, M. and O. Gürtler (2015), "The Optimality of Heterogeneous Tournaments," *Journal of Labor Economics*, 33, pp. 1007-1042.

Holmstrom, B. (1999), "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66, pp. 169-182.

Hunnes, A. (2012), "Testing the Role of Comparative Advantage and Learning in Wage and Promotion Dynamics," *International Journal of Manpower*, 33, pp. 556-582.

Jin, X. and M. Waldman (2017), "Lateral Moves, Promotions, and Task-Specific Human Capital: Theory and Evidence," Mimeo, Cornell University.

Kahn, L. (2013), "Asymmetric Information Between Employers," *American Economic Journal: Applied Economics*, 5, pp. 165-205.

Kahn, L. and F. Lange (2014), "Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures," *Review of Economic Studies*, 81, pp. 1575-1613.

Kauhanen, A. and S. Napari (2012), "Career and Wage Dynamics: Evidence from Linked Employer-Employee Data," *Research in Labor Economics*, 36, pp. 35-76.

Koch, A.K. and E. Peyrache (2009), "Aligning Ambition and Incentives," *Journal of Law, Economics, and Organization*," 27, pp. 655-688.

Krakel, M. and A. Schottner (2008), "Relative Performance Pay, Bonuses, and Job-Promotion Tournaments," IZA Discussion Paper No. 3702.

Kuhn, P. and F. Lozano (2008), "The Expanding Workweek? Understanding Trends in Long Work Hours Among U.S. Men, 1979-2006," *Journal of Labor Economics*, 26, pp. 311-343.

Lambert, R., D. Larcker, and K. Weigelt (1993), "The Structure of Organizational Incentives," *Administrative Science Quarterly*, 38, pp. 438-461.

Lazear, E. and P. Oyer (2013), "Personnel Economics, " in R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics*, Princeton University Press, pp. 479-519.

Lazear, E. and S. Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, pp. 841-864.

Lemiuex, T., W.B. MacLeod, and D. Parent (2009), "Performance Pay and Wage Inequality," *Quarterly Journal of Economics*, 124, pp. 1-49.

Lluis, S. (2005), "The Role of Comparative Advantage and Learning in Wage Dynamics and Intrafirm Mobility: Evidence from Germany," *Journal of Labor Economics*, 23, pp. 725-767.

Malcomson, J.M. (1984), "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92, pp. 486-507.

Milgrom, P. and J. Roberts (1988), "An Economic Approach to Influence Activities in Organizations," *American Journal of Sociology*, 94, pp. S154-S179.

Mukherjee, A. (2008), "Career Concerns, Matching, and Optimal Disclosure Policy," *International Economic Review*, 49, pp. 1211-1250.

Murphy, K.M. and R.H. Topel (1985), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics*, 3, pp. 370-379.

Murphy, K.M. and F. Welch (1990), "Empirical Age-Earnings Profiles," *Journal of Labor Economics*, 8, pp. 202-209.

Myerson, R.B. (1978), "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, pp. 73-80.

Nalebuff, B.J. and J.E. Stiglitz (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, 14, pp. 21-43.

O'Keefe, M., W.K. Viscusi, and R.J. Zeckhauser (1984), "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 2, pp. 27-56.

Ortin-Angel, P. and V. Salas-Fumas (1998), "Agency-Theory and Internal Labor-Market Explanations of Bonus Payments: Empirical Evidence from Spanish Firms," *Journal of Economics & Management Strategy*, 7, pp. 573-613.

Pinkston, J. (2009), "A Model of Asymmetric Learning with Testable Implications," *Review of Economic Studies*, 76, pp. 367-394.

Prendergast, C. (1999), "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, pp. 7-63.

Prendergast, C. and R.H. Topel (1996), "Favoritism in Organizations," *Journal of Political Economy*, 104, pp. 958-978.

Robinson, C. (1989), "The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models," *Journal of Political Economy*, 97, pp. 639-667.

Rosen, S. (1982), "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13, pp. 311-323.

Rubin, D. (1987), *Multiple Imputation for Nonresponse in Surveys*, vol. 17, John Wiley & Sons, Inc., New York.

Schönberg, U. (2007), "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 25, pp. 651-692.

Schulz, E., S. Chowdhury, and D. Van de Voort (2013), "Firm Productivity Moderated Link Between Human Capital and Compensation: The Significance of Task Specific Human Capital," *Human Resource Management*, 52, pp. 423-439.

Smeets, V. and F. Warzynski (2008), "Too Many Theories, Too Few Facts? What the Data Tell Us About the Link Between Span of Control, Compensation and Career Dynamics," *Labour Economics*, 15, pp. 688-704.

Waldman, M. (1984a), "Job Assignments, Signaling, and Efficiency," *Rand Journal of Economics*, 15, pp. 255-267.

Waldman, M. (1984b), "Worker Allocation, Hierarchies, and the Wage Distribution," *Review of Economic Studies*, 51, pp. 95-109.

Waldman, M. (2013a), "Theory and Evidence in Internal Labor Markets," in R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics*, Princeton University Press, pp. 520-574.

Waldman, M. (2013b), "Classic Promotion Tournaments Versus Market-Based Tournaments," *International Journal of Industrial Organization*, 31, pp. 198-210.

Waldman, M. and O. Zax (2016), "An Exploration of the Promotion Signaling Distortion," *Journal of Law, Economics, and Organization*, 32, pp. 119-149.

Wang, Y. (2013), *Essays on Promotion and Turnover*, Unpublished PhD Dissertation, Cornell University.

Zabojnik, J. (2012), "Promotion Tournaments in Market Equilibrium," *Economic Theory*, 51, pp. 213-240.

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

Table 1. *Summary Statistics (Single Firm Dataset)*

Job Level	N	Bonus payment	Salary	Bonus/Salary	Positive bonus	Promoted	Age	Tenure at level	Tenure at firm	Performance rating available	Performance Rating
1	4,977	550.73	35828.84	0.016	0.21	0.24	36.49	2.46	2.48	0.65	2.09
		(1767.7)	(8761.7)	(0.04)	(0.41)	(0.43)	(9.88)	(2.05)	(2.05)	(0.48)	(0.79)
2	5,633	1032.51	42144.56	0.025	0.28	0.18	38.13	2.92	4.81	0.75	2.07)
		(2549.0)	(8804.5)	(0.06)	(0.45)	(0.38)	(9.35)	(2.54)	(3.06)	(0.43)	(0.70)
3	6,434	1941.43	51600.68	0.038	0.40	0.11	38.94	3.24	6.25	0.74	1.83
		(3756.1)	(9606.7)	(0.06)	(0.49)	(0.31)	(8.98)	(2.65)	(3.61)	(0.44)	(0.68)
4	6,593	5148.74	78620.98	0.066	0.46	0.01	41.77	4.63	8.42	0.72	1.64
		(10506.2)	(19474.6)	(0.11)	(0.50)	(0.11)	(8.05)	(3.50)	(4.14)	(0.45)	(0.65)
All	23,637	2326.60	53562.94	0.038	0.34	0.13	39.02	3.38	5.59	0.72	1.88
		(6336.9)	(20952.7)	(0.08)	(0.48)	(0.33)	(9.22)	(2.90)	(3.96)	(0.45)	(0.72)

Note: This table displays means and standard deviations (reported in parentheses) of the key variables used in the empirical analysis. The unit of observation is worker-year. All statistics are computed from the whole working sample, whereas the statistics for 'Performance Rating' are computed using the subsample of observations for which performance ratings are available. Bonuses and salaries are reported in real 1998 dollars, and tenure variables are expressed in terms of years.

Table 2. *Summary Statistics (Multi-Firm Dataset)*

Job Level	N	Bonus payment	Salary	Bonus/Salary	Positive bonus	Promoted	Age	Tenure at level	Tenure at firm
1	60,435	66.02	2156.85	0.034	0.71	0.09	42.10	8.24	13.59
		(84.99)	(427.57)	(0.03)	(0.45)	(0.28)	(10.36)	(6.81)	(11.85)
2	185,681	123.39	2803.31	0.050	0.66	0.05	40.09	7.49	10.83
		(170.18)	(614.66)	(0.06)	(0.47)	(0.22)	(9.89)	(6.15)	(10.21)
3	137,571	322.08	3699.72	0.097	0.84	0.03	41.42	6.81	10.17
		(305.96)	(783.35)	(0.08)	(0.37)	(0.18)	(8.35)	(5.41)	(8.41)
All	383,687	185.59	3022.89	0.065	0.73	0.05	40.89	7.37	11.03
		(243.97)	(858.70)	(0.07)	(0.44)	(0.22)	(9.48)	(6.03)	(9.96)

Note: This table displays means and standard deviations (reported in parentheses) of the key variables used in the empirical analysis. The unit of observation is worker-year. Bonuses and salaries are reported in real euros, and tenure variables are expressed in terms of years.

Table 3. *Determinants of The Probability of Earning A Bonus (Single Firm Dataset)*

	A. Pooled OLS					B. Fixed-Effects				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Level=2	0.080***	0.082***	0.065***	0.067***	0.063***	0.206***	0.250***	0.235***	0.224***	0.232***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.015)	(0.017)	(0.018)	(0.020)	(0.025)
Level=3	0.206***	0.212***	0.188***	0.192***	0.160***	0.421***	0.510***	0.490***	0.465***	0.474***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.022)	(0.027)	(0.028)	(0.034)	(0.040)
Level=4	0.274***	0.286***	0.250***	0.262***	0.183***	0.452***	0.592***	0.566***	0.522***	0.497***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.030)	(0.039)	(0.041)	(0.052)	(0.061)
Promoted next year		0.056***	0.051***	0.053***	0.030**		0.089***	0.091***	0.080***	0.092***
		(0.009)	(0.009)	(0.009)	(0.012)		(0.013)	(0.013)	(0.013)	(0.017)
Average salary increase at current level			0.749***	0.671***				0.371***	0.336**	
			(0.071)	(0.072)				(0.133)	(0.132)	
Tenure at level=2				0.001	0.007				0.028***	0.023*
				(0.007)	(0.010)				(0.009)	(0.013)
Tenure at level=3				-0.018*	-0.011				0.029**	0.026
				(0.009)	(0.012)				(0.013)	(0.017)
Tenure at level=4				-0.023**	-0.021				0.013	0.025
				(0.011)	(0.014)				(0.017)	(0.021)
Tenure at level=5				-0.040***	-0.027*				-0.003	0.010
				(0.014)	(0.016)				(0.022)	(0.026)
Tenure at level>=6				-0.060***	-0.063***				-0.045*	-0.030
				(0.011)	(0.013)				(0.026)	(0.031)
Rating=2					-0.126***					-0.109***
					(0.008)					(0.013)
Rating=3					-0.282***					-0.218***
					(0.010)					(0.017)
N(worker-years)	23,637	23,637	23,637	23,637	17,000	23,637	23,637	23,637	23,637	17,000
Adjusted R^2	0.159	0.160	0.166	0.168	0.144	0.130	0.133	0.133	0.135	0.131
Log-likelihood	-13907	-13888	-13808	-13786	-10049	-7117	-7073	-7065	-7039	-4385
Test for joint significance (p-values for two-sided tests are reported)										
Job levels	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000
Tenure at level				<0.000	<0.000				<0.000	0.027
Rating					<0.000					<0.000

Note: This table displays the results of a linear probability model with the dependent variable taking on a value of one if the worker earns a bonus in the current year and zero if not. Panel A (columns 1-5) reports the results for the pooled OLS, and Panel B (columns 6-10) reports the results of the fixed-effects estimation. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator and clustered at the individual level. All pooled regressions include controls for the worker's age, race, gender, education and indicator variables for year, while the fixed-effects regressions include indicator variables for year. 'Level 1', 'Tenure at level 1' and 'Rating 1' are the omitted categories. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table 4. *Determinants of The Probability of Earning A Bonus (Multi-Firm Dataset)*

	A. Pooled OLS				B. Fixed-Effects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Level=2	-0.065***	-0.060***	-0.065***	-0.066***	-0.007	0.005	-0.002	0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Level=3	0.064***	0.072***	0.061***	0.059***	0.078***	0.100***	0.088***	0.100***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
Promoted next year		0.083***	0.082***	0.082***		0.034***	0.034***	0.029***
		(0.003)	(0.003)	(0.003)		(0.003)	(0.003)	(0.003)
Average salary increase at current level			0.027***	0.028***			0.033***	0.036***
			(0.002)	(0.002)			(0.001)	(0.002)
Tenure at level=2				0.037***				0.020***
				(0.002)				(0.002)
Tenure at level=3				0.038***				0.028***
				(0.003)				(0.002)
Tenure at level=4				0.038***				0.042***
				(0.003)				(0.003)
Tenure at level=5				0.030***				0.038***
				(0.003)				(0.003)
Tenure at level>=6				0.008**				0.041***
				(0.003)				(0.003)
N(worker-years)	383,687	383,687	383,687	383,687	383,687	383,687	383,687	383,687
Adjusted R^2	0.189	0.190	0.191	0.192	0.076	0.077	0.078	0.079
Log-likelihood	-192049	-191661	-191564	-191328	-29080	-28981	-28727	-28491
Test for joint significance (p-values for two-sided tests are reported)								
Job levels	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000
Tenure at level				<0.000				<0.000

Note: This table displays the results of a linear probability model with the dependent variable taking on a value of one if the worker earns a bonus in the current year and zero if not. Panel A (columns 1-4) reports the results for the pooled OLS, and Panel B (columns 5-8) reports the results of the fixed-effects estimation. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator and clustered at the individual level. All regressions include controls for the worker's age, gender, education, and indicator variables for year, firm size and industry. 'Level 1' and 'Tenure at level 1' are the omitted categories. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table 5. *Determinants of Bonus Payments (Single Firm Dataset)*

	OLS	OLS	FE	FE	FDIV	FDIV	NLIV	NLIV	NLIV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Level=2	0.048***	0.052***	0.206***	0.221***	0.148***	0.142***	0.280***	0.262**	0.238**
	(0.005)	(0.006)	(0.020)	(0.030)	(0.025)	(0.028)	(0.070)	(0.092)	(0.091)
Level=3	0.141***	0.138***	0.433***	0.446***	0.311***	0.273***	0.352**	0.260	0.573***
	(0.007)	(0.011)	(0.040)	(0.055)	(0.038)	(0.041)	(0.110)	(0.136)	(0.101)
Level=4	0.438***	0.333***	0.478***	0.436***	0.431***	0.355***	0.632***	0.421*	0.776***
	(0.018)	(0.019)	(0.074)	(0.109)	(0.061)	(0.068)	(0.132)	(0.182)	(0.147)
Age	0.002	-0.103**	-1.327***	-1.611***	0.582***	0.543***	0.656***	0.573***	0.599***
	(0.032)	(0.052)	(0.200)	(0.268)	(0.046)	(0.050)	(0.075)	(0.072)	(0.065)
Age^2	-5.798	6.523	56.980***	85.247***	-141.58***	-127.890***	-171.79***	-140.70***	-157.272***
	(3.856)	(5.458)	(19.261)	(26.405)	(12.36)	(13.534)	(20.20)	(18.45)	(18.630)
Tenure at level=2	0.034***	0.051***	0.100***	0.094***	0.053***	0.057***	0.069***	0.072***	0.068***
	(0.010)	(0.013)	(0.016)	(0.021)	(0.010)	(0.011)	(0.012)	(0.013)	(0.013)
Tenure at level=3	0.048***	0.066***	0.138***	0.104***	0.092***	0.092***	0.137***	0.129***	0.119***
	(0.012)	(0.015)	(0.021)	(0.027)	(0.016)	(0.018)	(0.021)	(0.021)	(0.021)
Tenure at level=4	0.052***	0.060***	0.150***	0.145***	0.122***	0.132***	0.189***	0.187***	0.171***
	(0.013)	(0.016)	(0.026)	(0.032)	(0.022)	(0.024)	(0.030)	(0.030)	(0.031)
Tenure at level=5	0.035**	0.029*	0.131***	0.110***	0.130***	0.133***	0.217***	0.197***	0.169***
	(0.015)	(0.017)	(0.035)	(0.042)	(0.028)	(0.030)	(0.040)	(0.040)	(0.040)
Tenure at level>=6	0.073***	0.060***	0.105**	0.093*	0.126**	0.095*	0.277***	0.195**	0.178**
	(0.019)	(0.018)	(0.041)	(0.052)	(0.038)	(0.041)	(0.062)	(0.062)	(0.059)
Promotion prize	-0.109***	-0.435***	-0.813***	-0.847***					
	(0.042)	(0.115)	(0.114)	(0.237)					
Rating=2		-0.121***		-0.111***		-0.073***		-0.105***	
		(0.014)		(0.022)		(0.016)		(0.022)	
Rating=3		-0.183***		-0.171***		-0.112***		-0.149***	
		(0.014)		(0.037)		(0.021)		(0.030)	
Unobserved Ability x Job Level									
Level=2							1.913***	1.615***	0.860**
							(0.312)	(0.446)	(0.332)
Level=3							1.163	0.858	1.239***
							(0.756)	(0.586)	(0.336)
Level=4							2.050*	1.821*	1.710***
							(0.880)	(0.915)	(0.442)

Performance Rating x Job Level									
Level=1									0.014
									(0.017)
Level=2									-0.010
									(0.022)
Level=3									-0.064***
									(0.019)
Level=4									-0.088***
									(0.023)
N(worker-years)	23,540	16,949	23,540	16,949	9,022	7,089	9,022	7,089	7,089
Test for joint significance (p-values for two-sided tests are reported)									
Job levels	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000
Age (at mean age)	0.17	0.19	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000
Tenure at level	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000
Additional tests (p-values for two-sided tests are reported)									
Level 2 = Level 3	<0.000	<0.000	<0.000	<0.000	<0.000	<0.000	0.66	0.99	0.002
Level 3 = Level 4	<0.000	<0.000	0.27	0.89	0.005	0.087	<0.000	0.051	0.062
Tenure 2 = Tenure 3	0.28	0.30	0.04	0.65	0.002	0.002	<0.000	<0.000	<0.000
Tenure 3 = Tenure 4	0.78	0.68	0.47	0.07	0.011	0.001	<0.000	<0.000	<0.000
Tenure 4 = Tenure 5	0.24	0.07	0.33	0.17	0.510	0.940	0.070	0.54	.87
Tenure 5 = Tenure 6	0.05	0.09	0.24	0.52	0.040	0.081	0.051	0.97	.73
Rating 2 = Rating 3		<0.000		0.01		0.005		0.027	

Note: The dependent variable is bonus in level in ten thousands of 1988 dollars. Variables 'Age' and 'Age^2' are rescaled by 100 and 100000, respectively. 'Level=1', 'Tenure at level=1' and 'Rating=1' are the omitted categories. To derive the expected promotion prize, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, the worker's age, education level, gender, race, tenure at the firm, job titles. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, race, job level, tenure at current level, and indicators for year. Columns 1-4 display the results of estimating equation (2), where columns 1 and 2 report the results for the pooled OLS, and columns 3 and 4 report the results for the fixed-effects estimation. In columns 5 and 6, GMM estimation is applied to equation (7). In these models, the instruments for the current job level assignment include the interaction between job level dummies and salary in t-2, a binary variable that takes a value of one if the worker is promoted at the end period t-2, and the salary increase from t-2 to t-1 (i.e., the difference between the worker's salary in period t-2 and that in period t-1). Also, interactions between these instruments and the tenure at the current level is used for efficiency reasons. In columns 7 and 8, GMM estimation is applied to equation (11). In these models, the lag of bonus payment is instrumented using the interaction between job level dummies and salary in t-2, a binary variable that takes a value of one if the worker is promoted at the end period t-2, and the salary increase from t-2 to t-1 (i.e., the difference between the worker's salary in period t-2 and that in period t-1). Also, interactions between these instruments and the current value of the exogenous variables are included. In column 9, the same set of instruments is used to apply GMM estimation to equation (13). Standard errors are reported in parentheses. For columns 1-4, they are obtained using the Huber-White sandwich estimator and clustered at the individual level. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table 6. *Determinants of Bonus Payments (Multi-Firm Dataset)*

	OLS (1)	FE (2)	FDIV (3)	NLIV (4)
Level=2	0.009***	-0.008***	0.042***	0.017***
	(0.001)	(0.001)	(0.002)	(0.003)
Level=3	0.087***	0.016***	0.081***	0.022***
	(0.001)	(0.001)	(0.001)	(0.004)
Age	0.006***	0.012***	-0.000	0.021*
	(0.000)	(0.000)	(0.010)	(0.009)
Age^2	-0.007***	-0.011***	0.002	-0.015
	(0.000)	(0.000)	(0.009)	(0.008)
Tenure at level=2	0.005***	0.001	0.004***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Tenure at level=3	0.007***	0.005***	0.008***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Tenure at level=4	0.011***	0.008***	0.017***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
Tenure at level=5	0.017***	0.010***	0.016***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)
Tenure at level>=6	0.008***	0.003***	0.014***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Promotion prize	-0.006***	-0.015***		
	(0.001)	(0.001)		
Unobserved Ability x Job Level				
Level=2				2.462***
				(0.230)
Level=3				2.818***
				(0.226)
N(worker-years)	383,687	383,687	259,403	259,403
Test for joint significance (p-values for two-sided tests are reported)				
Job levels	<0.000	<0.000	<0.000	<0.000
Age (at mean age)	<0.000	<0.000	0.980	0.014
Tenure at level	<0.000	<0.000	<0.000	<0.000
Additional tests (p-values for two-sided tests are reported				
Level 2 = Level 3	<0.000	<0.000	<0.000	0.137
Tenure 2 = Tenure 3	<0.000	<0.000	<0.000	<0.000

Tenure 3 = Tenure 4	<0.000	<0.000	<0.000	<0.000
Tenure 4 = Tenure 5	<0.000	0.001	0.550	0.010
Tenure 5 = Tenure 6	<0.000	<0.000	0.001	<0.000

Notes: The dependent variable is bonus in €10. The dependent variable is standardized to have mean zero and standard deviation of 1. Age and Age^2 are similarly standardized in columns 3 and 4. 'Level= 1', and 'Tenure at level=1' are the omitted categories. To derive the expected promotion prize, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, the worker's age, education level, gender, race, tenure at the firm, job titles. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, race, job level, tenure at current level, and indicators for year. Columns 1 (OLS) and 2 (Fixed effects) display the results of estimating equation (2). In column 3, GMM estimation is applied to equation (7). In this model, the instruments for the current job level assignment include the interaction between job level dummies and salary in t-2, a binary variable that takes a value of one if the worker is promoted at the end period t-2, and the salary increase from t-2 to t-1 (i.e., the difference between the worker's salary in period t-2 and that in period t-1). Also, interactions between these instruments and the tenure at the current level is used for efficiency reasons. In column 4, GMM estimation is applied to equation (11). In these models, the lag of bonus payment is instrumented using the interaction between job level dummies and salary in t-2, a binary variable that takes a value of one if the worker is promoted at the end period t-2, and the salary increase from t-2 to t-1 (i.e., the difference between the worker's salary in period t-2 and that in period t-1). Also, interactions between salary in t-2 and salary increase from t-2 to t-1 and the current period of the explanatory variables are used. Standard errors are reported in parentheses. For columns 1 and 2, they are obtained using the Huber-White sandwich estimator and clustered at the individual level. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table 7

The Trade-Off Between Bonus Payments and Promotion Prize with Parametric Estimates of Promotion Wage Increase (Single Firm Dataset)

	(1)	(2)	(3)	(4)	(5)	(6)
Promotion prize A	-0.180***	-0.187***	-0.204***	-0.219***	-0.453***	-0.486***
	(0.057)	(0.060)	(0.055)	(0.057)	(0.110)	(0.108)
Corrected standard error	[0.090]	[0.093]	[0.088]	[0.092]	[0.138]	[0.095]
p-value with corrected standard error	0.045	0.045	0.020	0.017	0.001	<0.001
N(worker-years)	20,468	20,468	20,468	20,468	15,469	15,469
Adjusted R^2	0.140	0.142	0.148	0.152	0.117	0.123
Log-likelihood	-206875	-206862	-206779	-206737	-156806	-156751
Promotion prize B	-0.201***	-0.216***	-0.201***	-0.224***	-0.511***	-0.536***
	(0.067)	(0.072)	(0.063)	(0.068)	(0.165)	(0.162)
Corrected standard error	[0.108]	[0.103]	[0.107]	[0.107]	[0.221]	[0.218]
p-value with corrected standard error	0.063	0.036	0.060	0.036	0.021	0.014
N(worker-years)	20,468	20,468	20,468	20,468	15,469	15,469
Adjusted R^2	0.140	0.142	0.148	0.152	0.117	0.123
Log-likelihood	-206875	-206861	-206782	-206739	-156808	-156755
Promotion prize C	-0.187***	-0.201***	-0.193***	-0.216***	-0.453***	-0.482***
	(0.059)	(0.063)	(0.056)	(0.060)	(0.137)	(0.133)
Corrected standard error	[0.098]	[0.109]	[0.101]	[0.106]	[0.168]	[0.161]
p-value with corrected standard error	0.055	0.065	0.055	0.042	0.007	0.003
N(worker-years)	20,468	20,468	20,468	20,468	15,469	15,469
Adjusted R^2	0.140	0.142	0.148	0.152	0.117	0.123
Log-likelihood	-206875	-206861	-206782	-206739	-156810	-156756
Explanatory Variables						
Tenure at current job level	No	Yes	No	Yes	No	Yes
Average salary increase at current job level	No	No	Yes	Yes	No	Yes
Performance ratings	No	No	No	No	Yes	Yes

Note: This table displays the results of estimating Equation (2), where we employed three different expected promotion prizes in estimation. To derive expected promotion prizes, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, and worker fixed-effects for ‘Promotion prize A’; the same variables (except for worker fixed-effects) and the worker’s age, education level, gender, race, tenure at the firm for ‘Promotion prize B’; and for ‘Promotion prize C’ we add job titles to the second set of control variables. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, race, job level, tenure at current level, and indicators for year. The dependent variable for all regressions reported in the table is bonus payments in 1988 dollars. All regressions include controls for the worker's age, race, gender, education level, job level and year in which the bonus is paid. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator, and they are corrected for the intraworker correlation. Corrected standard errors reported in brackets are obtained using a non-parametric bootstrap method, which is explained in the text. p-value refers to the two-tailed test for the significance of 'Promotion prize' using the corrected standard error. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table 8

The Trade-Off Between Bonus Payments and Promotion Prize with Parametric Estimates of Promotion Wage Increase (Multi-Firm Dataset)

	(1)	(2)	(3)	(4)
Promotion prize A	-1.494***	-1.553***	-1.477***	-1.537***
	(0.090)	(0.092)	(0.089)	(0.092)
Corrected standard error	[0.093]	[0.100]	[0.083]	[0.091]
p-value with corrected standard error	<0.001	<0.000	<0.000	<0.000
N(worker-years)	383,687	383,687	383,687	383,687
Adjusted R^2	0.386	0.388	0.389	0.391
Promotion prize B	-1.560***	-1.605***	-1.543***	-1.590***
	(0.092)	(0.095)	(0.091)	(0.094)
Corrected standard error	[0.083]	[0.086]	[0.091]	[0.097]
p-value with corrected standard error	<0.001	<0.000	<0.000	<0.000
N(worker-years)	383,687	383,687	383,687	383,687
Adjusted R^2	0.386	0.388	0.389	0.391
Promotion prize C	-1.493***	-1.543***	-1.478***	-1.531***
	(0.088)	(0.091)	(0.088)	(0.091)
Corrected standard error	[0.112]	[0.085]	[0.100]	[0.089]
p-value with corrected standard error	<0.001	<0.000	<0.000	<0.000
N(worker-years)	383,687	383,687	383,687	383,687
Adjusted R^2	0.386	0.388	0.389	0.391
Explanatory Variables				
Tenure at current job level	No	Yes	No	Yes
Average salary increase at current job level	No	No	Yes	Yes

Note: This table displays the results of estimating Equation (2), where we employed three different expected promotion prizes in estimation. To derive expected promotion prizes, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, and worker fixed-effects for ‘Promotion prize A’; the same variables (except for worker fixed-effects) and the worker’s age, education level, gender, tenure at the firm for ‘Promotion prize B’; and for ‘Promotion prize C’ we add job titles to the second set of control variables. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, job level, tenure at current level, and indicators for year. The dependent variable for all regressions reported in the table is bonus payments in real euros. All regressions include controls for the worker’s age, gender, education level, job level and year in which the bonus is paid. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator, and they are corrected for the intraworker correlation. Corrected standard errors reported in brackets are obtained using a non-parametric bootstrap method, which is explained in the text. p-value refers to the two-tailed test for the significance of 'Promotion prize' using the corrected standard error. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

Table Appendix B1. *Determinants of Bonus Payments (Single Firm Dataset)*

	OLS (1)	OLS (2)	FE (3)	FE (4)
Level=2	0.052***	0.035***	0.456***	0.350***
	(0.009)	(0.009)	(0.072)	(0.072)
Level=3	0.184***	0.139***	0.909***	0.736***
	(0.014)	(0.014)	(0.116)	(0.115)
Level=4	0.377***	0.242***	1.202***	0.856***
	(0.016)	(0.017)	(0.182)	(0.179)
Age	-0.094	-0.175**	-3.053***	-3.011***
	(0.068)	(0.072)	(0.496)	(0.545)
Age^2	4.911	15.217**	132.658***	172.175***
	(7.362)	(7.642)	(49.872)	(55.840)
Tenure at level=2	0.033*	0.060***	0.136***	0.159***
	(0.019)	(0.022)	(0.036)	(0.039)
Tenure at level=3	0.015	0.035*	0.202***	0.168***
	(0.019)	(0.020)	(0.047)	(0.044)
Tenure at level=4	0.053**	0.091***	0.309***	0.252***
	(0.023)	(0.025)	(0.056)	(0.055)
Tenure at level=5	-0.012	-0.010	0.325***	0.222***
	(0.023)	(0.023)	(0.069)	(0.066)
Tenure at level>=6	0.025	0.015	0.369***	0.250***
	(0.023)	(0.021)	(0.087)	(0.083)
Promotion prize	-0.257**	-0.521***	-0.608**	-1.017***
	(0.102)	(0.117)	(0.290)	(0.297)
Rating=2		-0.129***		-0.100***
		(0.016)		(0.034)
Rating=3		-0.204***		-0.148***
		(0.017)		(0.050)
N(worker-years)	8,977	7,066	8,977	7,066
Adjusted R^2	0.071	0.083	0.540	0.592
Log-likelihood	-7554	-5345	-4401	-2487
Test for joint significance (p-values for two-sided tests are reported)				
Job levels	<0.000	<0.000	<0.000	<0.000
Age (at mean age)	0.46	0.041	0.002	<0.000
Tenure at level	0.032	<0.000	<0.000	<0.000

Additional tests (p-values for two-sided tests are reported)				
Level 2 =Level 3	<0.000	<0.000	0.66	<0.000
Level 3 = Level 4	<0.000	<0.000	<0.000	0.15
Tenure 2 = Tenure 3	0.29	0.202	0.033	0.06
Tenure 3 = Tenure 4	0.055	0.017	0.019	0.027
Tenure 4 = Tenure 5	0.003	<0.000	0.67	0.49
Tenure 5 = Tenure 6	0.102	0.224	0.245	0.42
Rating 2 = Rating 3		<0.000		0.21

Note: The dependent variable is bonus in level in ten thousands of 1988 dollars. Variables ‘Age’ and ‘Age2’ are rescaled by 100 and 100000, respectively. ‘Level=1’, ‘Tenure at level=1’ and ‘Rating=1’ are the omitted categories. To derive the expected promotion prize, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, the worker’s age, education level, gender, race, tenure at the firm, job titles. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, race, job level, tenure at current level, and indicators for year. Columns 1-4 display the results of estimating equation (2), where columns 1 and 2 report the results for the pooled OLS, and columns 3 and 4 report the results for the fixed-effects estimation. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator and clustered at the individual level. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level

Table Appendix B2. *Determinants of Bonus Payments (Multi-Firm Dataset)*

	OLS (1)	FE (2)
Level=2	0.009***	-0.005***
	(0.001)	(0.001)
Level=3	0.087***	0.018***
	(0.001)	(0.002)
Age	0.006***	0.019***
	(0.000)	(0.001)
Age^2	-0.007***	-0.017***
	(0.000)	(0.001)
Tenure at level=2	0.003***	-0.003***
	(0.001)	(0.001)
Tenure at level=3	0.010***	0.003***
	(0.001)	(0.001)
Tenure at level=4	0.014***	0.005***
	(0.001)	(0.001)
Tenure at level=5	0.011***	0.002*
	(0.001)	(0.001)
Tenure at level>=6	0.011***	-0.001
	(0.001)	(0.001)
Promotion prize	-0.006***	-0.014***
	(0.001)	(0.001)
N(worker-years)	259403	259403
Adjusted R^2	0.343	0.134
Log-likelihood	-241181	-344269
Test for joint significance (p-values for two-sided tests are reported)		
Job levels	<0.000	<0.000
Age (at mean age)	<0.000	<0.000
Tenure at level	<0.000	<0.000
Additional tests (p-values for two-sided tests are reported)		
Level 2 =Level 3	<0.000	<0.000
Tenure 2 = Tenure 3	<0.000	<0.000
Tenure 3 = Tenure 4	<0.000	0.116
Tenure 4 = Tenure 5	0.002	<0.000
Tenure 5 = Tenure 6	0.610	<0.000

Notes: The dependent variable is bonus in €10. The dependent variable is standardized to have mean zero and standard deviation of 1. ‘Level= 1’, and ‘Tenure at level=1’ are the omitted categories. To derive the expected promotion prize, we estimate equations (16) and (17). In estimating equation (16), we employ the set of explanatory variables including job level, tenure at current level, year, the worker’s age, education level, gender, race, tenure at the firm, job titles. The set of explanatory variables used in the estimation of equation (17) includes the worker age, education, gender, race, job level, tenure at current level, and indicators for year. Standard errors reported in parentheses are obtained using the Huber-White sandwich estimator and clustered at the individual level. <0.000 indicates that the corresponding p-value is smaller than 0.0005. ***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.