

Interpreting Intra-firm Wage Differentials using Tournament Models*

Matthew Shum
Dept. of Economics
Johns Hopkins University

Preliminary: June 24, 2005
Comments welcome: please do not quote

Abstract

We consider estimation of tournament models of intra-firm labor markets. We require only data on intra-firm employment and wages, and do not require observations of individual worker productivity. We illustrate our procedure using a dataset of intra-firm wage differentials for a sample of major retail chains in the U.S. We find that effort levels are generally higher at higher strata of employment within a firm, but that only a fraction (typically less than 50%) of the wage differential directly compensates workers for higher effort levels, implying that over half of the differentials arise purely to maintain incentives at lower rungs of the company.

1 Introduction

Wage differentials within retail chains, even at the lowest levels, can be quite large: for example, full-time sales staff at a major clothing retail chain are paid roughly \$10,000 per annum on average, but store managers earn around \$21,000, over twice that amount. Further up the hierarchy, district managers make on average over \$38,000 (in 1986 dollars).¹

*This draft is preliminary and may contain errors. Please do not quote. I am grateful to Joe Harrington for providing me access to the data used in this analysis, and for useful comments. I also thank Nick Hill, Ben Klemens, Philippe Marcoul, Joanne Roberts, Chris Ruebeck, Nadia Soboleva, Ayako Suzuki, and seminar participants at Northwestern University for useful comments. Please contact me at Dept. of Economics, Johns Hopkins University, Baltimore, MD 21218; e-mail mshum@jhu.edu.

¹These data are drawn from the National Retail Federation Specialty Store Wage and Benefit Survey, which will be described below. For confidentiality reasons, the names of the stores used in this study cannot

Why do these differentials arise? A standard explanation for these differentials is that employees at higher levels of a firm are paid more, because they work harder, or are more productive. An alternative explanation is provided in the tournament literature, where wages at the top of a hierarchy must be kept high in order to provide incentives for workers, even in low levels of the hierarchy, to exert effort.

Despite the sizable theoretical literature on tournament models (see McLaughlin (1989) for a survey), empirical work related to these models is limited. Previous empirical work on tournament models have mainly focused on testing the predictions of these models. This includes papers on executive compensation (cf. Main, O’Reilly, and Wade (1993), Eriksson (1999)), professional sports (eg. Ehrenberg and Bognanno (1990a), (1990b), Bronars and Oettinger (2001)), and agricultural (poultry) production (Knoeber and Thurman (1994)).

Most closely related to the present paper is a series of papers by Ferrall (1996), (1997), who estimates structural models of internal labor markets within, respectively, law firms and engineering firms. The main difference between those papers and the present paper lies in the empirical approach, which to large extent is dictated by the datasets used in the different studies. Relatedly, Ferrall and Smith (1999) develop and estimate a structural tournament-like model of sports championship series.

One feature of much of the existing empirical work is that the researcher observes productivity measures at the individual worker level, which are difficult to obtain (or unavailable) in practice. In this paper, we develop methodologies for estimating tournament models which require only firm-level information on employment and wages (at different hierarchical levels within a firm), and do not require observation of workers’ productivity levels. By exploiting the equilibrium restrictions of the elimination tournament model, we derive estimates of model unobservables — including workers’ equilibrium effort levels — which are consistent with the observed wage data. Hence, the aim of this paper is not to test the tournament theory (as in the previous empirical work), but to use it as a guide to obtain estimates of the “structural” parameters of the tournament model.

We illustrate our methodologies using a unique dataset of wages and employment levels within a number of large retail chains (including many retailers found in typical shopping malls). Our empirical analysis focuses on the lower hierarchical levels (i.e., the sales staff, assistant store manager, and store manager positions) of these firms, in which internal promotion appears to be the most common way to advance and, therefore, the tournament

be mentioned in the paper. By definition, full-time sales staff must work at least 30 hours a week.

model appears most appropriate. Our results suggest that only a small fraction – typically less than 50% – of the observed wage differentials directly compensates workers for higher effort at higher levels of the hierarchy, implying that over half of the differentials arise purely to maintain incentives at lower rungs of the company.

Since performance measures are not readily available for many industries and firms, the methodology developed in this paper may also be useful for analyzing other industries or sectors. In the next section, we present a store-level tournament model, based on the model in Rosen (1986), that we will employ in this paper. In Section 3, we develop and discuss the estimation methodology, and Section 4 describes the data. We present empirical results in Sections 5 and 6. We conclude in Section 7.

2 Economic model

The tournament view of a firm’s internal labor market differs in important respects from the efficiency wage literature on labor contracts, which likewise focuses on contracts as a means to provide incentives to workers to provide effort. The non-tournament efficiency wage literature has focused on “absolute” compensation schemes, in which each worker is paid according to how her observed performance measures against some objective, absolute benchmark. As long as a worker’s observed performance depends (even stochastically) on her effort or productivity, these non-tournament compensation schemes ultimately need not generate intra-firm wage differentials if effort or productivity is unchanging across different hierarchical levels within the firm.

This is not the case with tournaments, which are a form of “relative” compensation schemes, whereby a given worker is paid (or promoted) depending on her performance relative to her co-workers. In an elimination tournament, workers must exert effort and be productive even at low levels of the hierarchy in order to remain in contention for the larger prizes which are available at higher levels of the firm.² As Rosen (1986) points out, intra-firm wage differentials can arise even when workers exert identical effort levels at each stratum of the firm, because the pay differentials at higher levels of the hierarchy motivate effort exertion

²Another strand of the tournament literature has focused on rank-order tournaments, in which workers are paid according to their relative performance in identical tasks (cf. Lazear and Rosen (1981), Green and Stokey (1983), and Holmstrom (1982)). The goal of rank-order tournaments is to encourage high effort in a homogeneous task in the presence of common (across all workers) unobserved productivity shocks, which differs from the goal of an elimination tournament, which is generally to provide incentives for continued (and perhaps higher) effort levels at higher hierarchical strata of the company. This paper focuses solely on the elimination-type of tournaments, since they appear to offer a more appropriate description of the hierarchical structure of the retail firms which we study.

at lower levels of the hierarchy. In this paper, we use a unique dataset containing salaries and employment levels across hierarchical levels within a number of major retail firms in order to estimate a tournament model based on Rosen’s model. We are able to estimate the equilibrium effort levels consistent with the observed data and the theoretical tournament model.

There are various informational reasons that a firm would employ a tournament scheme to determine its payments to its workers, when the effort levels of individual workers are not directly observable.³ First, the noisy measure of effort which the firm observes may be difficult to interpret cardinally, which motivates the use of compensation schemes based only on ordinal rankings of these measures (such as tournaments). Second, there may be common unobservable productivity shocks across workers. Third, workers may be heterogeneous, and tournaments would be a way of identifying and rewarding the “best” worker. In this paper, we assume that workers are homogeneous and, hence, abstract away from the third point above. However, our model makes no explicit assumptions regarding the information structure, and accommodates both the first and second features above.

We introduce a simple framework, based on the elimination tournament model of Rosen (1986), which captures to some extent the promotional possibilities within the hierarchy of a given firm. In this model, each worker’s career within a firm transpires as a progression through a tournament, in which workers compete against each other at each hierarchical level of the firm, with the winners at each level surviving to compete at higher levels. Rosen studies the relation between the wage structure within a firm and workers’ incentives to exert effort at each level of the firm. The goal of this paper is to uncover the parameters of workers’ preferences and equilibrium effort levels which are consistent with the observed wage sequences from a number of retail chains.

A given firm has $S + 1$ hierarchical levels, indexed by s , with $s = 0$ corresponding to the highest level, and $s = S$ corresponding to the entry level. W_1, \dots, W_{S+1} denote the payoffs (wages) at each level of the firm. (Note our indexing convention, whereby W_{s+1} is the wage that the “losers” at stage s make.) Hence, W_{S+1} is the salary earned by the employees at the lowest level in the firm hierarchy, and it can be interpreted as a “reservation wage” for all the workers in our model.

³Tournament (or compensation) design is not too relevant without asymmetric information between workers and firms. In a symmetric information situation, firms can just compensate workers directly for their effort costs: even when information is imperfect, as long as it is symmetric (as in the matching literature), the use of wages to provide incentives to workers is not a relevant issue. We return to this issue below, when discussing our results.

Let n_0, \dots, n_S denote the number of workers at each level of the firm: hence, the total number of workers from level s who get promoted up to level $s - 1$ are $\sum_{s=0}^{s-1} n_s \equiv m_{s-1}$. Note that, by this definition, $m_{s-1} < m_s$, for $s = 0, \dots, S$.

2.1 Competition

The strengths and limitations of the dataset dictate to some extent the empirical model that we employ in this paper. While the dataset contains detailed information on the salaries and number of employees at different hierarchical levels within a number of retail chains, it contains no measures of output or performance. Within each level of the firm, we specify a model of competition which seems especially relevant for the retail environment. Within level s , we assume that the firm divides the m_s contenders into L_s equal-sized subgroups, each consisting of m_s/L_s workers (abstracting away from integer issues). A tournament is played among the members of each subgroup, with the $m_{s-1}/L_s \geq 1$ best performers selected to advance to the next level $s - 1$. Let

$$f_s \equiv m_s/L_s, \quad g_s \equiv m_{s-1}/L_s$$

denote, respectively, the number of contenders and winners per subgroup. In the retail application below, a subgroup is naturally interpreted as an individual store.⁴

We assume that all workers are homogeneous, and focus on a symmetric pure strategy Nash equilibrium in which each worker exerts the identical effort level x_s^* (where $x > 0$) at level s .⁵ An agent who exerts an effort level \bar{x} during level s while all of the rivals in her subgroup exert the equilibrium level of effort x_s^* progresses with probability

$$P_s(\bar{x}; x_s^*) = \frac{h(\bar{x}) + (g_s - 1) * h(x_s^*)}{h(\bar{x}) + (f_s - 1) * h(x_s^*)}. \quad (1)$$

In the above, $h(\cdot)$ is a link function translating individual effort levels into the probability of advancement.⁶ The $P_s(\dots)$ function captures, in reduced-form, the procedure whereby

⁴In Rosen's (1986) paper, and in most sports tournaments (eg. tennis tournaments, the soccer World Cup), $f_s = 2$ for $s = 1, \dots, S$, so that the number of subgroups at each level $L_s = \frac{1}{2}m_s$.

⁵It is difficult to accommodate worker heterogeneity since our dataset contains no information on employees, besides their average salaries. However, we note the caveat that by assuming homogeneity of employees, we abstract away from one important theoretical explanation of tournaments: to pick out the "best" candidate among a field of heterogeneous contestants.

⁶The odds-ratio parameterization of the advancement probability conditional on effort follows Rosen. For the $f_s = 2$ case considered by Rosen, if $h(x) = \exp(x)$, then the advancement probability (1) is a binary logit probability. This form of the advancement probability function can be justified by a model where the worker i with the higher productivity y_{is} advance out of stage s , and the productivity measure y_{is} is equal to the effort x_{is} plus an additive random noise term which follows the type I extreme value distribution. This structural interpretation of the odds-ratio parameterization (1) no longer applies when $f_s \neq 2$.

the firm selects winners at each stage of the tournament, and is induced ultimately by the information structure of the game (i.e., what signals of effort the firm observes). Since our dataset includes no measures of productivity for any employee, we avoid more detailed modeling of the information structure, and adopt the reduced-form advancement probability given in Eq. (1).⁷

In the symmetric equilibrium, all employees exert identical levels of effort, so that $\bar{x} = x_s^*$ and the probability of surviving level s does not depend on the equilibrium effort level x_s^* :

$$p_s^* \equiv P_s(x_s^*, x_s^*) = \frac{g_s}{f_s} = \frac{m_{s-1}}{m_s}. \quad (2)$$

While the functional form for the advancement probability in Eq. (1) is not arbitrary, as remarked above, the form of the equilibrium winning probability (2) obtains very generally, requiring only that any g_s -subset of the f_s contestants in stage s of the tournament are chosen to advance with equal probability. This is a natural assumption because, in equilibrium, every contestant expends an identical level of effort.

Note that *any* symmetric equilibrium in which all m_s contestants exert identical levels of effort (including zero effort) will yield the same winning probability m_{s-1}/m_s in equilibrium. Therefore, the specific values of f_s and g_s matter only insofar as it affects the players' incentives, and therefore the effort levels that they choose. Indeed, as we will see below, the toughness of competition (as parameterized by f_s and g_s) has a crucial effect on the amount of effort exerted in equilibrium.⁸

2.2 Equilibrium

The equilibrium sequence of effort levels $\{x_s^* : s = 1, \dots, S\}$ is determined by a dynamic optimization problem.⁹ Let V_s denote the (equilibrium) value of progressing to (and potentially beyond) level s . If a given worker chooses effort level \bar{x} at level s , her value V_s is implicitly defined via the Bellman equation¹⁰

$$V_s = \max_{\bar{x}} \{P_s(\bar{x}; x_s^*)V_{s-1} + (1 - P_s(\bar{x}; x_s^*))W_{s+1} - c(\bar{x})\} \quad (3)$$

⁷However, the exact functional form for this advancement probability is not arbitrary, and can imply some restrictions on the information structure. See Ferrall (1996), pp. 814–815, for an example where the form of the advancement function $P_s(\dots)$ is explicitly derived from the information structure of the game.

⁸Additionally, as Rosen notes, the symmetric equilibrium has a prisoner's dilemma quality: every worker would be better off if nobody exerted any effort (since the equilibrium winning probability P_s is the same no matter how much effort is exerted), but this is not an equilibrium.

⁹With $S + 1$ hierarchical levels, there are only S stages to the tournament, because there is no more competition at the top $S = 0$ stage.

¹⁰Implicitly, we are assuming a discount rate equal to 1.

where $c(\cdot)$ is the cost of effort function. In the above display, the first term within the curly brackets denotes the worker's expected payoff from advancing to the next ($s - 1$ -th) round, while the second term is the payoff from "losing" in stage s and obtaining the wage W_{s+1} . At level s , a given worker chooses an effort level \bar{x} to maximize the right-hand side of (3).

In the symmetric equilibrium, all workers expend identical effort levels x_s^* in level s ; this effort level must satisfy the following first-order condition:

$$P_{s,1}(x_s^*; x_s^*) (V_{s-1} - W_{s+1}) - c'(x_s^*) = 0 \quad (4)$$

where $P_{s,1}(\dots)$ denotes the derivative of $P_s(\dots)$ with respect to the first argument. Given the odds-ratio parameterization (1) of the $P_s(\dots)$ function, in equilibrium

$$p_{s,1}^* \equiv \left. \frac{\partial P_s(x; x_s^*)}{\partial x} \right|_{x_s^*} = \frac{h'(x_s^*)}{h(x_s^*)} \frac{1 - p_s^*}{f_s}. \quad (5)$$

By substituting Eq. (5) into the first-order condition (4), we obtain

$$\frac{h'(x_s^*)}{h(x_s^*)} \frac{1 - p_s^*}{f_s} (V_{s-1} - W_{s+1}) = c'(x_s^*). \quad (6)$$

The above equation is the main equation which characterizes equilibrium effort levels in our tournament game. By completely differentiating it, we can derive several comparative statics. We make the assumptions that $h'(x) > 0$, $c'(x) > 0$, $c''(x) \geq 0$, for all effort levels x . First, if $h''(x_s^*) \leq 0$, then $\frac{dx_s^*}{dp_s^*} < 0$: equilibrium effort levels are smaller when the equilibrium probability of advancement to the next stage is higher. (However, if the h function is convex, the sign is ambiguous.)

Second, we examine the effect of an increase in the number of contestants f_s , while holding the advancement probability fixed (i.e., by always adjusting the number of winners $g_s = f_s \cdot p_s^*$, for some pre-specified level of p_s^*). If (as above), we assume the concavity of the h function, then $\frac{dx_s^*}{df_s} < 0$: equilibrium effort levels are smaller when the number of contestants increases.¹¹ One interpretation of this finding is that an increase in the "toughness of competition" (as measured by f_s) actually dilutes equilibrium incentives to provide effort.

This point is illustrated in Fig. (1), where we have graphed pairs of best response curves corresponding to different values of f and g , values of (respectively) the contenders and winners from the lowest (sales staff) level of a tournament played at the San Francisco-area stores of a major clothing retailer (these data area given in Table 2, which we discuss

¹¹In making this calculation, we assume that any change in f_s leaves p_{s-1}^* unaffected.

below). In the solid lines, we graph the best response curves corresponding to the case where f (the number of contending sales staff) is 10.25, and the number of winners g (those who “win” the sales floor tournament) is 4.38, the actual observed values. The lines marked with circles are the best response curve for the case where both f and g are halved: clearly, as the comparative statics predict, the reduction in the number of competitors has raised workers’ incentives to exert effort, and the equilibrium effort levels (measured in money units) double from about \$1500 to \$3000. In contrast, if we double the number of contenders, effort levels decrease by about half, from about \$1500 to \$800 units (as indicated by the intersection of the third set of best-response curves, marked in the crossed lines). An increase in the toughness of competition (as measured by an increase in the number of contenders) actually reduces equilibrium incentives to provide effort: this is because, as stated above, when there are more competitors, the marginal effect of additional effort on the winning probability is lowered.

[Figure 1 about here.]

The first-order condition (6) also forms the basis of our empirical methodology, which has as its goal the estimation the the equilibrium effort levels x_1^*, \dots, x_S^* as well as (to the extent possible, as we will be precise about later) the functions $h(\cdot)$ and $c(\cdot)$. We describe this methodology next.

3 Empirical strategy: two approaches

The first-order condition (6) can be rewritten as

$$\mu(x_s^*) \frac{1 - p_s^*}{f_s} (V_{s-1} - W_{s+1}) = c(x_s^*), \quad (7)$$

where

$$\mu(x) \equiv \frac{h'(x)}{h(x)} / \frac{c'(x)}{c(x)}. \quad (8)$$

In matrix notation, this is

$$\begin{pmatrix} \frac{\mu(x_1^*)}{f_1}(1 - p_1^*) & 0 & \cdots & 0 \\ 0 & \frac{\mu(x_2^*)}{f_2}(1 - p_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\mu(x_S^*)}{f_S}(1 - p_S^*) \end{pmatrix} \left[\begin{pmatrix} V_0 \\ V_1 \\ \vdots \\ V_{S-1} \end{pmatrix} - \begin{pmatrix} W_2 \\ W_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] = \begin{pmatrix} c(x_1^*) \\ c(x_2^*) \\ \vdots \\ c(x_S^*) \end{pmatrix}. \quad (9)$$

This can be conveniently used in order to derive a recursive formulation for V_s : by plugging Eq. (7) into Eq. (3), we obtain

$$V_s = \beta_s V_{s-1} + (1 - \beta_s) W_{s+1}, \quad s = 1, \dots, S \quad (10)$$

where

$$\beta_s \equiv p_s^* - \mu(x_s^*) \frac{1 - p_s^*}{f_s}.$$

Using the initial condition $V_0 = W_1$, we can solve Eq. (10) forward to derive

$$V_1 = \beta_1 W_1 + (1 - \beta_1) W_2$$

$$V_2 = \beta_2 \beta_1 W_1 + \beta_2 (1 - \beta_1) W_2 + (1 - \beta_2) W_3$$

\vdots

$$V_S = \beta_1 \dots \beta_S W_1 + (1 - \beta_1) \beta_2 \dots \beta_S W_2 + (1 - \beta_2) \beta_3 \dots \beta_S W_3 + \dots + (1 - \beta_S) W_{S+1}$$

or

$$\begin{pmatrix} V_0 \\ V_1 \\ V_2 \\ \vdots \\ V_S \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \beta_1 & (1 - \beta_1) & \dots & 0 & 0 \\ \beta_1 \beta_2 & \beta_2 (1 - \beta_1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_1 \dots \beta_S & (1 - \beta_1) \beta_2 \dots \beta_S & \dots & (1 - \beta_{S-1}) \beta_S & (1 - \beta_S) \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_S \\ W_{S+1} \end{pmatrix}. \quad (11)$$

Substituting (11) into (9), we obtain

$$\begin{pmatrix} \frac{\mu(x_1^*)}{f_1} (1 - p_1^*) & 0 & \dots & 0 \\ 0 & \frac{\mu(x_2^*)}{f_2} (1 - p_2^*) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\mu(x_S^*)}{f_S} (1 - p_S^*) \end{pmatrix} \left[\begin{pmatrix} 1 & 0 & \dots & 0 \\ \beta_1 & (1 - \beta_1) & \dots & 0 \\ \beta_1 \beta_2 & \beta_2 (1 - \beta_1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \dots \beta_{S-1} & (1 - \beta_1) \beta_2 \dots \beta_{S-1} & \dots & (1 - \beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_S \end{pmatrix} - \begin{pmatrix} W_2 \\ W_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] = \begin{pmatrix} c(x_1^*) \\ c(x_2^*) \\ \vdots \\ c(x_S^*) \end{pmatrix}. \quad (12)$$

The system of equations (12) gives us, for each firm, S equations but $2S$ unknowns $c(x_1^*), \dots, c(x_S^*), \mu(x_1^*), \dots, \mu(x_S^*)$. In order to proceed, we must make some additional assumptions. Next, we describe the two approaches we employ in this paper to circumvent this fundamental underidentification issue.

3.1 First approach

In the first approach, we address the under-identification issues by making several functional form assumptions. Since the units of effort are arbitrary, we normalize these units so that each unit of effort leads to a disutility of one dollar: this implies that the cost of effort function $c(x) = x$. However, even with this assumption, the expressions for $\mu(x)$ do not simplify.

Therefore, we use a further assumption that $h(x) = x$. However, this functional form assumption is *not* a normalization, because it restricts $\mu(x) = 1, \forall x$. Given this assumption, the system simplifies to

$$\begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_S^* \end{pmatrix} = \begin{pmatrix} \frac{1}{f_1}(1-p_1^*) & 0 & \cdots & 0 \\ 0 & \frac{1}{f_2}(1-p_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{f_S}(1-p_S^*) \end{pmatrix} \left[\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \beta_1 & (1-\beta_1) & \cdots & 0 \\ \beta_1\beta_2 & \beta_2(1-\beta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \cdots \beta_{S-1} & (1-\beta_1)\beta_2 \cdots \beta_{S-1} & \cdots & (1-\beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_S \end{pmatrix} - \begin{pmatrix} W_2 \\ W_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] \quad (13)$$

which gives S equations in the S unknowns x_1^*, \dots, x_S^* . We can therefore solve for x_1^*, \dots, x_{S-1}^* from observed data on wages W_1, \dots, W_{S+1} and intra-firm employment levels m_0, \dots, m_S .

3.2 Second approach

The main benefit of the first approach is that we are able to recover distinctive effort levels for each set of (firm-, geographic region-, and year-) observations. However, this is done

at the cost of making a potentially restrictive functional form assumption regarding the link function $h(\cdot)$. In the second approach, we model the $h(\cdot)$ function more flexibly, but substitute instead the assumption that for a given firm, the effort levels x_1^*, \dots, x_S^* , as well as the functional form of the $h(\cdot)$ function, are constant over both time and geographic locations. We continue to assume that $c(x) = x$, as this is nothing but a normalization.

However, we introduce randomness into the model by allowing the wages W_0, \dots, W_S to be observed with error. One may be justified in assuming the presence of measurement error, because the observed wages are obtained by surveys, where respondents report the average salaries earned in each hierarchical level.

In particular, we assume that W_{ismt} , the observed wage for firm i , strata s , location m , and year t , is equal to the actual (but unobserved) wage W_{ismt}^* perturbed with additive measurement error ϵ_{ismt} :

$$W_{ismt} = W_{ismt}^* + \epsilon_{ismt}, \quad s = 1, \dots, S \quad (14)$$

where ϵ_{ismt} is a mean zero measurement error assumed independent across s , m , and t for a given firm i . Note that we assume that W_{iS+1mt} , the wage at the lowest stratum of the company, is not observed with error. The reason for this will be noted below.

We estimate the parameters x_1^*, \dots, x_S^* , as well as the parameters of the $h(\cdot)$ function, separately for each firm i . (In what follows, we drop the firm i subscript for convenience.) For each firm, we will estimate these parameters by method of moments. In order to derive the estimating equations, we combine Eqs. (11), (5), and (14) to obtain

$$\left[\begin{pmatrix} \frac{h'(x_1^*)/h(x_1^*)}{f_1}(1-p_1^*) & 0 & \dots & 0 \\ 0 & \frac{h'(x_2^*)/h(x_2^*)}{f_2}(1-p_2^*) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{h'(x_S^*)/h(x_S^*)}{f_S}(1-p_S^*) \end{pmatrix}^* \right. \\ \left. \left[\begin{pmatrix} 1 & 0 & \dots & 0 \\ \beta_1 & (1-\beta_1) & \dots & 0 \\ \beta_1\beta_2 & \beta_2(1-\beta_1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \dots \beta_{S-1} & (1-\beta_1)\beta_2 \dots \beta_{S-1} & \dots & (1-\beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 - \epsilon_1 \\ W_2 - \epsilon_2 \\ \vdots \\ W_S - \epsilon_S \end{pmatrix} - \begin{pmatrix} W_2 - \epsilon_2 \\ W_3 - \epsilon_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] \right] = \begin{pmatrix} c'(x_1^*) \\ c'(x_2^*) \\ \vdots \\ c'(x_S^*) \end{pmatrix} \quad (15)$$

or, in shorthand,

$$\begin{aligned} \mathbf{A} \left[\mathbf{B} \left(\vec{W}_{1:S} - \vec{\epsilon}_{1:S} \right) - \left(\vec{W}_{2:S+1} - \vec{\epsilon}_{2:S} | 0 \right) \right] &= \vec{c}' \Rightarrow \\ \vec{\epsilon} &= -\tilde{\mathbf{B}}^{-1} \left[\mathbf{A}^{-1} \vec{c}' - \mathbf{B} \vec{W}_{1:S} + \vec{W}_{2:S+1} \right] \end{aligned} \quad (16)$$

where $\tilde{\mathbf{B}}$ is the matrix \mathbf{B} minus a $S \times S$ matrix with ones in the $(i, i + 1)$, $i = 1, \dots, S - 1$ spots and zeros everywhere else. (In the first display, $\vec{\epsilon}_{2:S} | 0$ denotes the S -vector where the first $S - 1$ elements are $\epsilon_2, \dots, \epsilon_S$ and the S -th element is a zero.) Because the system of equations (15) is only S -dimensional, we cannot accommodate an additional measurement error in the wage W_{S+1} .¹²

We assume a power (constant-elasticity) specification for the h function:

$$h(x) = x^\gamma \quad (17)$$

so that a total of $S + 1$ parameters— $\theta \equiv (x_1^*, \dots, x_S^*, \gamma)$ are estimated.¹³ For this specification of $h(\cdot)$, γ parameterizes the responsiveness of the advancement probability P_s to an individual's effort level. Hence, in a setting where effort is observable with noise, it is reasonable to interpret a larger value of γ as implying that the observations of effort are less noisy.

With this assumption, the system of first-order conditions (15) reduces to

¹²Since this is ad-hoc, we also obtained results where we allowed W_{S+1} to be observed with error, but assumed instead that the observed W_1 (the district manager salary) contained no measurement error. Overall, the magnitude of the results remained quite stable.

¹³Note that if we were to assume that $c(x) = x^\alpha$, the first-order condition in Eq. (7) would reduce to

$$\frac{\gamma}{\alpha} \frac{1 - p_s}{f_s} (V_{s-1} - W_{s+1}) = x^\alpha$$

which is the same as a model where $c(x) = x$ but $h(x) = x^{\gamma/\alpha}$. In this sense, then, the maintained assumption that $c(x) = x$ is just a normalization.

$$\left[\begin{pmatrix} \frac{\gamma}{x_1^* f_1} (1 - p_1^*) & 0 & \cdots & 0 \\ 0 & \frac{\gamma}{x_2^* f_2} (1 - p_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\gamma}{x_S^* f_S} (1 - p_S^*) \end{pmatrix} * \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \beta_1 & (1 - \beta_1) & \cdots & 0 \\ \beta_1 \beta_2 & \beta_2 (1 - \beta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \cdots \beta_{S-1} & (1 - \beta_1) \beta_2 \cdots \beta_{S-1} & \cdots & (1 - \beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 - \epsilon_1 \\ W_2 - \epsilon_2 \\ \vdots \\ W_S - \epsilon_S \end{pmatrix} - \begin{pmatrix} W_2 - \epsilon_2 \\ W_3 - \epsilon_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (18)$$

and $\beta_s = p_s - \gamma \frac{1-p_s}{f_s}$ for $s = 1, \dots, S$.

Furthermore, the power-function specification also implies that the comparative statics discussed at the end of Section 2.2 can be unambiguously signed, regardless of the curvature of the $h(\cdot)$ function, and yield that both $\frac{dx_s^*}{dp_s^*} < 0$ and $\frac{dx_s^*}{df_s} |_{p_s^* \text{ fixed}} < 0$. That is, both an increase in the advancement probability, and an increase in the level of competition (without an accompanying increase in p_s^*) lead to lower effort levels in equilibrium.

In addition, with the power function parameterization, we also obtain that $\frac{dx_s^*}{d\gamma} > 0$: the equilibrium effort levels should be increasing in γ , the power parameters attached to the link function $h(\cdot)$. Fig. 2 illustrates how changes in γ affects equilibrium incentives to exert effort. The solid lines are a pair of best-response curves corresponding to $\gamma = 2.21$ (the value we estimate for retailer I, using the second approach), When γ is increased to 2.5, implying more responsiveness in the probability of winning to worker effort, we see that the circle-marked best response curves intersect at a higher point, implying that equilibrium effort levels rise, by around 15%. When γ is decreased increased to 2.0 (illustrated by the best-response curves marked with crosses), the equilibrium effort levels decrease by around 15%.

For each firm i , the population moment conditions exploited in the estimation is

$$E[\epsilon Z] = 0 \quad (19)$$

where Z is an M -vector of instruments. The sample analog of the above is

$$\mathbf{m}_{ST}(\theta) \equiv \begin{bmatrix} \frac{1}{ST} \sum_{t=1}^T \sum_{s=1}^S \epsilon_{st}(\theta) Z_{s1t} = 0 \\ \vdots \\ \frac{1}{ST} \sum_{t=1}^T \sum_{s=1}^S \epsilon_{st}(\theta) Z_{sMt} = 0 \end{bmatrix}$$

where the dependence of ϵ_{st} on the parameters θ emphasizes the fact that, at each value of θ , the ϵ 's are obtained as residuals, via Eq. (16). Let $M \geq S + 2$ be the total number of moments conditions employed in estimating θ . We seek the minimizer of the quadratic form

$$\theta_{MST} \equiv \operatorname{argmin}_{\theta} \mathbf{m}_{MST}(\theta)' W \mathbf{m}_{MST}(\theta)$$

and our estimator has the limiting distribution (as T goes to infinity)

$$\sqrt{ST} (\theta_{MST} - \theta_0) \xrightarrow{d} N(0, (J' \Omega J)^{-1} J' \Omega V \Omega J (J' \Omega J)^{-1})$$

where

$$J = E_0 \frac{\partial \mathbf{m}(\theta_0)}{\partial \theta_0}$$

$$V = \operatorname{Var}_0 \mathbf{m}(\theta_0) = E_0 \mathbf{m}(\theta_0) \mathbf{m}(\theta_0)'$$

and $\mathbf{m}(\theta)$ denotes the $S + 2$ -vector of moment conditions, and Ω is a $M \times M$ weighting matrix. In practice, we use a two-step GMM procedure in which an estimate of the optimal weighting matrix $\Omega = V^{-1}$ issued in the second step, so that the limiting variance of our estimator reduces to $(J' \Omega J)^{-1}$.

Second-order optimality conditions In both estimation approaches given above, we assume that the first-order condition (4) characterizes the optimal effort levels chosen by the workers. However, second-order conditions should also hold at the optimal effort levels, given the tournament parameters. In our empirical work, therefore, we check each estimated set of effort levels x_1^*, \dots, x_S^* to ensure that they satisfy the second-order condition which corresponds to the first-order condition in Eq. (4):

$$\hat{c}'(\hat{x}_s^*) * \left[\frac{\hat{h}''(\hat{x}_s^*)}{\hat{h}'(\hat{x}_s^*)} - 2 \frac{\hat{h}'(\hat{x}_s^*)}{f_s * \hat{h}(\hat{x}_s^*)} \right] - \hat{c}''(\hat{x}_s^*) < 0 \quad (20)$$

for $s = 1, \dots, S$, and the hats ($\hat{\cdot}$) denoting estimated quantities. For the constant-elasticity (power) functional form assumption on the $h(\cdot)$ function used in the second approach, the

second-order conditions reduce to

$$\frac{\hat{\gamma} - 1}{\hat{x}_s^*} - 2 \frac{\hat{\gamma}}{f_s \hat{x}_s^*} < 0.$$

In what follows, we report only the empirical results which satisfy these second-order conditions. Since we do not impose these conditions directly in obtaining our estimates, they constitute, informally, specification checks on the model.

4 Data

We illustrate the methodologies developed above using data on wage differentials and employment levels at a number of large US retail chains. Most of these retailers commonly have locations in shopping malls and centers in the suburban US.

The dataset is drawn from the Specialty Store Wage and Benefit Survey performed by the National Retail Federation (NRF), for the years 1997-1999. This survey contains information on the number of employees and average annual salary for employees at various levels of the store hierarchy, for a number of large retail chains. For confidentiality reasons, we are not able to identify the chains by name, but refer to them by letters (see Table 1 for a list of the 14 chains considered in the empirical exercise).

The data are aggregated up to the (retail chain–geographic area) level, so that we cannot distinguish between different stores within the same chain and geographic area. Therefore, in our analysis, we essentially treat all the stores within the geographic area as identical stores. We focus only on four levels of employment within each chain: Sales Staff ($S = 3$), Assistant Store Manager ($S = 2$), and Store Manager ($S = 1$), with the District Manager ($S = 0$) position taken as the prize in the tournament. The reason we focus only on these levels of the firm is because only in the lower levels is internal promotion — a maintained assumption of the tournament model — likely to be prevalent. At levels higher than the district manager, it is much more likely that the chain may hire from the outside.

For these three levels of competition, we define the number of subgroups separately for each (firm-geographic location) pair. At the sales staff and assistant store manager levels, we take $L_2 = L_3 = n_1$, the number of store managers in each geographic region. That is, we assume that competition takes place within stores, and that the number of stores that a retail chain operates in each geographic region is equal to the total number of store managers employed in each region. At the store manager level, we assume that $L_1 = n_0$, the total number of district managers in the geographic region. That is, we assume that

the firm divides up the stores within each geographic region in a number of equal-sized districts, each headed by a single district manager. (Throughout, we ignore integer issues for convenience.)

In Table 1 we give summary statistics on the wage and tournament parameters (number of contenders and winners in each subgroup, at each of the three stages considered) for the retail chains which we will consider in our study, averaged across geographic locations and years. (For fulltime Sales Staff, salary is calculated as hourly wage*40 hours*50 weeks.) There are large variations across stores both in wages as well as the intensity of competition, so that we perform estimation on a store-by-store basis. Across most retailers, the assistant store manger/store manager salary gap is at least as large, and often larger, than the sales staff/assistant store manager gap. Note also that g_2 , the number of “winners” in the second stage (the competition among assistant store managers) is always slightly more than one; this is because, even though competition takes place within each store in stages 2 and 3, each store must promote slightly more than one winner, in order to produce the required number of employees to cover both the store manager and district manager positions.¹⁴

Twelve of the fourteen chains studied in this analysis report that the compensation of their sales-staff are “straight salary” meaning that less than 25% of their compensation is commission-based. The two exceptions are retailer B in 1998 (the only year for which this is reported for this retailer), and retailer C in 1998¹⁵, which reported that the sales staff were compensated in a fashion such that over 25% of the compensation arose from commissions. Note that both retailers B and C are footwear retailers (retailer B sells athletic shoes, while retailer C sells dress shoes).

Remarks Before presenting the results, we note that the model presented above is stylized, and abstracts away from several potentially important features of the retailers that we consider. First, the model does not accommodate voluntary attrition (quits) out of the tournament which, in practice, is likely to be very common, especially among the sales staff of a retail chain. Indeed, in our dataset, the median turnover rate (defined as the number of terminations divided by the number of workers employed at a given level) across firms is 14%, 23%, 33%, and 38% at the district manager, store manager, assistant store manager, and (full-time) sales staff stages, respectively. However, this feature need not affect our

¹⁴Note that we do not consider part-time sales staff in this paper, and assume that these employees do not participate in the tournament. Typically, the number of part-time sales staff outnumber full-time sales staff by a ratio of 5:1 or 6:1 in most retail chains.

¹⁵In 1997, this retailer reported that its sales-staff compensation was “straight salary”.

analysis since we interpret the data at hand as representing the *desired* level of employment at each hierarchy. Therefore, we allow for sales staff to quit, as long as they are replaced by other workers who “take their place” in the tournament (and as long as the assistant store managers and store managers are chosen from among the sales staff, not from the outside). Nevertheless, in the Appendix we consider an expanded model which allows employees to leave the firm if they receive higher outside wage offers, and estimate this model for the subset of retailers for which we observe enough turnover data.

Second, we assume that all employees enter the hierarchy by the lowest (sales staff) stage. In practice, however, it may be possible to be directly hired as an assistant store manager, or store manager. We are unable to assess how important this is among the retailers in our dataset; hence, we reiterate that these considerations encourage interpreting the results presented below with some caution.

[Table 1 about here.]

5 Empirical results

5.1 First approach

As an example, Table 2 presents estimates using the first approach (where we set both $c(x) = h(x) = x$), using data from retailer I, for the year 1999. Estimated effort levels are presented for each geographic area separately. There are no “standard errors” for the values of x_1^*, \dots, x_S^* obtained by this procedure, since the obtained values are just transformations of the observed data. Throughout, we define wages without including any bonuses.¹⁶

[Table 2 about here.]

In order to gauge how important the competition assumptions (cf. Section 2.1 above) are to the estimates of equilibrium effort levels, we re-estimated under the alternative assumption that in a given stage s , the retail chain selects the m_{s-1} top performers among the m_s contenders for advancement, so that $L_s = 1$, and $f_s = m_s$ and $g_s = m_{s-1}$ (this maintains the same promotional probabilities as in Table 2). As we would expect, effort levels are substantially lower under the alternative assumptions (we do not report these results for

¹⁶We have re-estimated all the models defining wages to include bonuses, and obtained qualitatively similar results to those reported here. We do not report these additional results for the sake of brevity.

the sake of brevity). For example, the store-level competition results in Table 2 imply that sales staff at Boston-area stores expend the equivalent of \$76.89 in effort, but under the alternative assumptions the effort level falls to only \$30.92. This is because when the link function $h(\cdot)$ is linear, equilibrium effort levels should be decreasing in f_s , the number of competitors, holding the advancement probability constant.

Table 4 presents the average estimates for effort levels for each retailer, where the average is taken across all (year, geographic region) pairs observed in the dataset for that particular chain.¹⁷

[Table 3 about here.]

Two features are noticeable. First, higher effort levels are exerted at higher levels of the firm, implying that at least part of the observed wage differentials across the store manager, assistant store manager, and sales staff stages can be justified for effort reasons alone. Second, for a majority of the firms, there is a larger gap in absolute effort between the assistant store manager and store manager stages than between the sale staff and assistant store manager stages. This tends to mimic the relative sizes of the wage differentials, as given in Table 1.

5.2 Results: second approach

Table 5 presents estimates of effort levels, and the parameters of the cost function $c(\cdot)$ and $h(\cdot)$ function, using the second, GMM-based approach.¹⁸ In estimation, we employ seven moment conditions to estimate the five parameters in the model. The instruments which we use for a given observation of firm i , stratum s , location m , and year t are (i) n_{ismt} , the number of workers in the firm at this location, year and level; and (ii) $w_{ism't}$, the wages of workers in the same level and during the same year, but at different locations $m' \neq m$.

[Table 4 about here.]

Generally, the estimates are reasonably precisely estimated. As with the first-approach estimates reported before, we check that, at the estimated quantities, the second-order

¹⁷As discussed above, for each set of estimates, we checked that the second-order conditions in Eqs. (20). Since these conditions are not directly imposed in obtaining the estimates, their satisfaction can suggest (informally) that the tournament model is not too inappropriate in explaining the observed wage differentials and employment levels at each level of the retail chains.

¹⁸We varied the starting values in obtaining our estimates, to ensure that the reported estimates are reasonably robust and stable.

conditions (in Eq. (20)) are satisfied at the estimated effort levels. As a formal specification check, the GMM J -statistic (and the associated p -value under the null of correctly-specified moment conditions) is given in the last column. For all 14 retail chains, we could not reject the null hypothesis (at reasonable significance levels) that the moment conditions are indeed satisfied at the estimated values.

The estimates for the $h(\cdot)$ function parameter, γ , indicate that this function is convex across all retailers¹⁹, ranging from a low value of 1.34 for retailer J, to 3.86 for retailer H. While the order of magnitude of the estimated effort levels is the same as for the first approach estimates, note that the second approach estimates indicate generally higher effort levels than the first approach estimates. This is because we estimate $\gamma > 1$ across all fourteen retailers (whereas $\gamma = 1$ was imposed in obtaining the first approach results): as we discussed above, equilibrium effort levels increase when γ increases.

Furthermore, a simple correlation table for the estimated parameters shows that both x_1 and x_2 (but not x_3) are negatively correlated with γ across retailers. In addition, effort levels are generally higher at higher levels of the company. The sole exception is retailer J, for which a pronounced drop in effort occurs between the ASM stage (where effort costs of \$3341.8 are expended) and the SM stage (where effort costs of \$1912.7 – about a 40% drop – are expended). Retailer J is the sole eyewear retailer, and arguably this retailer is the one where sales staff are required to have the most training, as their duties include fitting and ordering eyewear for customers.

Given the variation in estimated effort levels and γ 's across retailers, we ran regressions of these estimated quantities on retailer characteristics. For each retailer, we created five dummy variables, for line of business and also location and size characteristics (which are defined and summarized in Table 6). Subsequently, we performed regressions of the estimated γ 's and effort levels on these dummy variables.

[Table 5 about here.]

The regression results are reported in Table 7. For γ , the parameter in the $h(\cdot)$ function linking effort to the probability of winning, we see from the second column of Table 7 that *DFOOT* enters positively and significantly, while *DLOC* enters negatively, but only marginally significantly. To the extent that a larger γ may be due to less noisy observations

¹⁹Given footnote 13 above, this finding can also be interpreted to imply that the $h(\cdot)$ function is more convex than the cost of effort function across all 14 retailers.

of effort, this result implies that effort is better measured in footwear retailers: this may be because selling shoes requires more direct customer services (eg. measuring customers and fetching the right-sized shoes, etc.) than, say, selling clothing. Curiously, this effect is not due to the fact that two of the four footwear retailers (B and C) in our dataset pay sales-staff based on commissions: we see from Table 5 that the γ 's for retailers B and C (which are 2.23 and 2.49, respectively) are lower than the estimated γ for retailer H (3.86), another footwear retailer which doesn't use commission-based pay to compensate sales staff.

In the third column of Table 7 we ran a regression of the pooled estimated effort levels on the retailer characteristics, as well as dummy variables for the three different hierarchical levels. As expected, the hierarchical dummies enter significantly, reflected the fact that higher effort is exerted at higher levels of the hierarchy. None of the retailer characteristics enter significantly.

[Table 6 about here.]

[Figure 2 about here.]

6 Do wage differentials reward greater effort levels?

With results in hand, we return to the question posed at the beginning of the paper: how much of the observed wage differentials arise to compensate workers for exerting higher effort, and how much to provide incentives? In a perfect-information, perfectly-competitive setting, intra-firm wage differentials between two positions should just compensate employees for their effort cost differentials across the two positions. In a tournament setting, this need not be true, since wage differentials between levels i and $i + 1$ must also serve to give incentives for more effort at lower levels $s > i$ of the company. Indeed, Rosen (1986) presents an example in which workers' effort levels are the same in every stage of the company, where positive wage differentials arise purely to provide incentives at lower stages of the tournament.

Using the results obtained above, we can directly measure how much of the observed intra-firm wage differentials between stages i and $i + 1$, $w_i - w_{i+1}$, directly rewards higher effort expenditures, $c(x_i) - c(x_{i+1})$. We introduce the notation $\delta_{st} \equiv 100 * \left(\frac{c_s - c_t}{w_s - w_t} \right)$, $s < t$, which denotes the ratio of effort to wage differential between stages s and t . In Table 8 we present the values for δ_{12} (between the assistant store manager and store manager stages) and δ_{23} (between the sales staff and assistant store manager stages) implied by our model

estimates. Note that some of these estimates are quite imprecise, with standard errors which are the same order of magnitude as (and often exceeding) the point estimates. Therefore we interpret these results with some caution. We focus on the second approach results (reported in the rightmost column of Table 8) in the following discussion.

[Table 7 about here.]

All the estimated δ 's are less than 100. The median value across all retailers and hierarchical levels is 20.3, and the average value is 27.3.

Since wages at a given stage i can provide incentives for effort only at stages $s > i$ prior to i , this implies that, for most chains, wages at the assistant store manager stage are an important source of incentives for effort provision on the sales floor and, similarly, wages at the store manager level also compensate for effort exerted in earlier stages. Indeed, for several stores, the estimated δ is *negative*. This occurs only when the effort differential between stages is estimated to be negative (because the wage differentials are never negative), and implies that the wage differential exists completely to compensate effort at previous stages.

6.1 How optimal are tournaments?

The results from the previous section suggest that a large proportion — typically over 50% — of the observed wage differentials arise to provide workers incentives to exert effort. The need for these incentives would only arise when effort is directly unobservable, and therefore non-contractible. While tournaments are attractive in these asymmetric information situations because they are robust to a variety of imperfections in the effort measurement system (including noisy effort measures, availability only of ordinal (but not cardinal) measures of effort, as well as common observables in the noisy effort measures taken across individuals) they cannot overcome the fundamental difficulties in effort measurement, and so attain only second-best outcomes. In order to gauge how well the second-best tournaments are performing for the retail chains which we study, we compare each retail chain's observed wage bill under the tournament setting to its wage bill under a first-best scheme in which workers' effort levels are directly observed, and therefore compensation based directly on a worker's effort.

At each hierarchical level s , the first-best wage (assuming firms have complete bargaining

power), for a given effort level x_s in stage s , is given by

$$W_s^{FB} = W^R + x_s$$

where W^R is some reservation wage (assumed constant across all stages s), and x_s is the effort level at stage s , in money units.

In calculating the first-best wage bill for each store, we assume that the firm's desired effort levels at each stage correspond to the effort levels estimated before. Furthermore, we set W^R for each store to be equal to $W_{S+1} - x_S$, the sales staff salary less the cost of effort for each salesperson.²⁰

These wage bill results are given in Table 9. We see that, indeed, for all of the stores, and across both approaches, the first-best wage bill is lower than the observed wage bill (which we interpret throughout this paper as being generated from a second-best tournament). However, the percentage differences between the two wage bills (reported in columns 3 and 6 of Table 9 for, respectively, the first and second approach estimates) are not unreasonably high, ranging from 11% for Retailer H, up to 43% for Retailer J (using the second approach results). Therefore, while tournaments are second-best, in some cases the firms are not doing much worse using the tournaments, compared with the first-best scenario.²¹

[Table 8 about here.]

7 Conclusions

In this paper, we have shown that, we can extract a lot of information from data on intra-firm wage differentials consistent with equilibrium in tournament models. We have developed methodologies for recovering equilibrium effort levels from data on intra-firm employment levels and salaries. The estimates suggest that effort levels are generally higher at higher strata of employment within a firm, but that only a small fraction of the wage differential directly compensates workers for higher effort levels: at the estimated effort levels, we find that typically less than 50% of the observed wage differentials are for rewarding higher effort

²⁰We note that we cannot compute the entire first-best wage bill for all stages of the tournament (from $S=0$ to $S=3$) because we cannot estimate the equilibrium effort level exerted by district managers (which will determine the wage in a first-best setting).

²¹Ideally, one would like to compare the tournament wage bill to the wage bill from another second-best incentive pay scheme. However, this would require additional assumptions regarding the performance measures that the firms observe, upon which the workers' wages would depend. This is difficult because we observe no information on these matters in the data.

levels at higher levels of the corporations, implying that over half of the differentials arise purely to maintain incentives at lower rungs of the company. We reiterate here that the model we use is stylized, and so these results should be taken with some caution.

There are also more general extensions of the current work. One important implicit assumption made in this paper is that the tournament framework is correct, and no attempt has been made to test the tournament framework versus alternative models of the data generating process for the observed wage data. Moreover, we have assumed here that workers are homogeneous within a firm, across all hierarchical levels. Hence, an interesting extension is to allow for workers to be endowed with varying levels of “talent” (as in Rosen’s paper). In this way, more talented workers will tend to be selected into the upper echelons of the company so that, conditional on surviving, equilibrium effort levels may actually be lower at the upper echelons. This motivates the interesting question as to whether high-end managerial salaries reward inherent talent, or increased effort. We wish to address these issues in future work, but we envision that it can be quite challenging without data on workers’ output or performance.

References

- BRONARS, S., AND G. OETTINGER (2001): "Effort, Risk-Taking, and Participation in Tournaments: Evidence from Professional Golf," Manuscript, University of Texas at Austin.
- EHRENBERG, R., AND M. BOGNANNO (1990a): "Do tournaments have incentive effects?," *Journal of Political Economy*, 98, 1307–1324.
- (1990b): "The incentive effects of tournaments revisited: Evidence from the European PGA Tour," *Industrial and Labor Relations Review*, 43, S74–S88.
- ERIKSSON, T. (1999): "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data," *Journal of Labor Economics*, 17, 262–280.
- FERRALL, C. (1996): "Promotions and Incentives in Partnerships: Evidence from Major U.S. Law Firms," *Canadian Journal of Economics*, 29, 811–827.
- (1997): "Empirical Analysis of Occupational Hierarchies," *Journal of Human Resources*, 34, 1–34.
- FERRALL, C., AND A. SMITH (1999): "A Sequential Game Model of Sports Championship Series: Theory and Estimation," *Review of Economics and Statistics*, 81, 704–719.
- GREEN, J., AND N. STOKEY (1983): "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 91, 349–364.
- HOLMSTROM, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324–340.
- KNOEBER, C., AND W. THURMAN (1994): "Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production," *Journal of Labor Economics*, 12, 155–179.
- LAZEAR, E., AND S. ROSEN (1981): "Rank-order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89, 841–864.
- MAIN, B., C. O'REILLY, AND J. WADE (1993): "Top executive pay: Tournament or teamwork?," *Journal of Labor Economics*, 12, 606–628.
- MCLAUGHLIN, K. (1989): "Aspects of Tournament Models: A Survey," in *Research in Labor Economics*, ed. by R. Ehrenberg. JAI Press, Inc.
- ROSEN, S. (1986): "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76, 701–715.

A Extension: accommodate turnover rates

In this section, we describe how the model estimated in the paper could be extended to allow workers to quit the firm if they receive a wage offer which exceeds their current salary. In our modification of the model, we allow employees to leave the firm once they find their place in the firm (i.e., after they “lose” and remain in some stage s). We do not allow workers to quit while they are still “active” in the tournament. In addition, we maintain the assumption that the firm can only hire workers from the outside at the lowest level of the firm (level S), and that positions at levels $s < S$ can only be filled by promoting workers from lower stages.

In the dataset, we observe turnover rates ρ_0, \dots, ρ_S , where ρ_s is defined as the ratio of the total workers terminated in stage s divided by the total number of workers employed at stage s (which includes both the terminated and non-terminated workers).²² These turnover rates are observed at the firm and year level, but only at the national level (i.e., not broken down by geographic locations).

Let F_s , $s = 0, \dots, S$ denote the CDF of outside wages for employees in stage s . We can interpret the observed turnover rates as

$$\rho_s = 1 - F_s(W_{s+1}), \quad s = 0, \dots, S. \quad (21)$$

That is, the observed turnover rate at stage s is interpreted as the probability of obtaining an outside wage offer exceeding the stage s salary, which is W_{s+1} . The workers’ Bellman equation, for stage s , is

$$V_s = \max_x \{p_s(x; x_s^*)V_{s-1} + (1 - p_s(x; x_s^*)) E_{R_s} \max(W_{s+1}, R_s)\} \quad (22)$$

where R_s denotes the outside wage offer for a stage s worker, and $R_s \sim F_s$. Obviously,

$$E \max(W_{s+1}, R_s) = F_s(W_{s+1})W_{s+1} + (1 - F_s(W_{s+1})) E[R_s | R_s > W_{s+1}].$$

Let $\tilde{W}_{s+1} \equiv E \max(W_{s+1}, R_s)$.

In order to estimate this amended model, we need to make additional assumption on the outside wage distributions F_0, \dots, F_S . In the following, we assume that each of the wage distributions is uniform:

$$R_s \sim \mathcal{U}[0, \bar{\theta}_s]$$

²²Obviously, firms do not fire workers in our model. All terminations arise because the worker receive a higher outside wage offer and quit.

where $\bar{\theta}_s$, $s = 0, \dots, S$ are unknown parameters. With this distributional assumption:

$$\begin{aligned} F_s(W_{s+1}) &= \frac{W_{s+1}}{\bar{\theta}_s} = 1 - \rho_s \\ \Rightarrow \bar{\theta}_s &= \frac{W_{s+1}}{1 - \rho_s}, \quad s = 0, \dots, S \end{aligned} \quad (23)$$

$$\begin{aligned} E[R_s | R_s > W_{s+1}] &= \frac{1}{2} (W_{s+1} + \bar{\theta}_s) \\ &= \frac{1}{2} W_{s+1} \frac{2 - \rho_s}{1 - \rho_s} \end{aligned}$$

$$\tilde{W}_{s+1} = W_{s+1} * \left[(1 - \rho_s) + \frac{1}{2} \frac{\rho_s(2 - \rho_s)}{(1 - \rho_s)} \right].$$

Hence, after plugging in these items into Eq. (22), we can estimate as before.

For the first approach, given observations of ρ_0, \dots, ρ_S , we can back out $\bar{\theta}_0, \dots, \bar{\theta}_S$ using Eq. (23). Then we can construct $\tilde{W}_1, \dots, \tilde{W}_{S+1}$ and then estimate x_1^*, \dots, x_S^* using the system of equations (13), substituting in \tilde{W}_s in place of W_s , for $s = 1, \dots, S + 1$.

For the second approach, we again assume that the observed wages are contaminated by additive measurement error. With this assumption:²³

$$\begin{aligned} \tilde{W}_s &= (W_s - \epsilon_s) * \left[(1 - \rho_s) + \frac{1}{2} \frac{\rho_s(2 - \rho_s)}{(1 - \rho_s)} \right], \quad s = 1, \dots, S \\ \tilde{W}_{S+1} &= W_{S+1} * \left[(1 - \rho_{S+1}) + \frac{1}{2} \frac{\rho_{S+1}(2 - \rho_{S+1})}{(1 - \rho_{S+1})} \right]. \end{aligned}$$

Let $\Psi_{1:S}$ denote the $S \times S$ diagonal matrix with $(1 - \rho_s) + \frac{1}{2} \frac{\rho_s(2 - \rho_s)}{(1 - \rho_s)}$ in the s -th diagonal position. Then, for this case, the estimating equation corresponding to Eq. (16) in the main text is

$$\begin{aligned} \mathbf{A} \left[\mathbf{B} \Psi_{0:S-1} \left(\vec{W}_{1:S} - \vec{\epsilon}_{1:S} \right) - \Psi_{1:S} \left(\vec{W}_{2:S+1} - \vec{\epsilon}_{2:S} | 0 \right) \right] &= \vec{c}' \Rightarrow \\ \vec{\epsilon}' &= -\tilde{\mathbf{B}}^{-1} \left[\mathbf{A}^{-1} \vec{c}' - \mathbf{B} \Psi_{0:S-1} \vec{W}_{1:S} + \Psi_{1:S} \vec{W}_{2:S+1} \right] \end{aligned}$$

where

$$\tilde{\mathbf{B}} \equiv \mathbf{B} \Psi_{0:S-1} - \tilde{\Psi}_{2:S-1}$$

²³As above, we need to assume that one of the wages – in this case, W_{S+1} – is not contaminated by measurement error.

and $\tilde{\Psi}_{2:S-1}$ denotes the $\Psi_{1:S-1}$ matrix bordered at the bottom and the left with, respectively, a row and column of zeros.

We continue to assume that the observed employment levels for each firm, year, and geographic location are the desired employment levels for the firm. For both the first and second approaches, the presence of turnover implies that firms must hire, and promote, more workers in order to achieve the desired employment levels at each stage of the firm. Hence, we need to reconstruct our measures of the number of competitors at each stage. The L_s 's (number of subgroups) stay the same. We must redefine the number of "losers" and "contenders" in each stage as (for $s = 0, \dots, S$):

$$\tilde{n}_s = \frac{n_s}{1 - \rho_s}$$

$$\tilde{m}_s = \sum_{s=0}^s \frac{n_s}{1 - \rho_s}.$$

Results using the second approach for the four retailers for which we were able to obtain turnover rates at the sales staff, assistant store manager, store manager, and district manager levels, for at least a single year, are reported in Table 10. Noticeably, the J -test specification checks have substantially lower p -values for this model, relative to the results which are in the main text, which are obtained from a model which does not accommodate turnover.²⁴

However, the qualitative implications of the results are similar to the above results. The results in Tables 11 and 12 are the turnover model analogs to those reported in, respectively, Tables 8 and 9 for the no-turnover model.

[Table 9 about here.]

[Table 10 about here.]

[Table 11 about here.]

²⁴However, the second-order conditions in Eq. (20) continue to hold for these estimates.

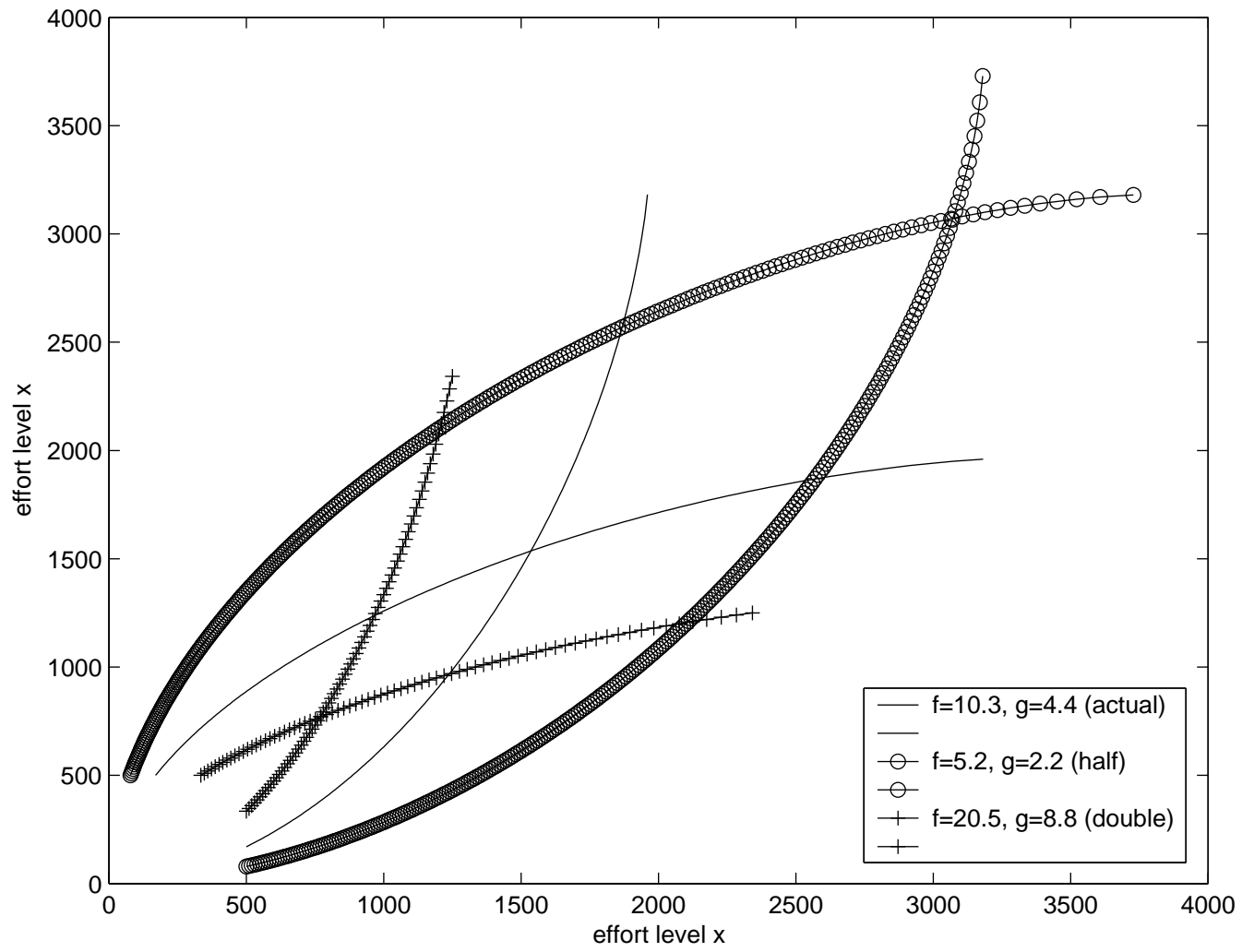


Figure 1: Effects of changes in toughness of competition

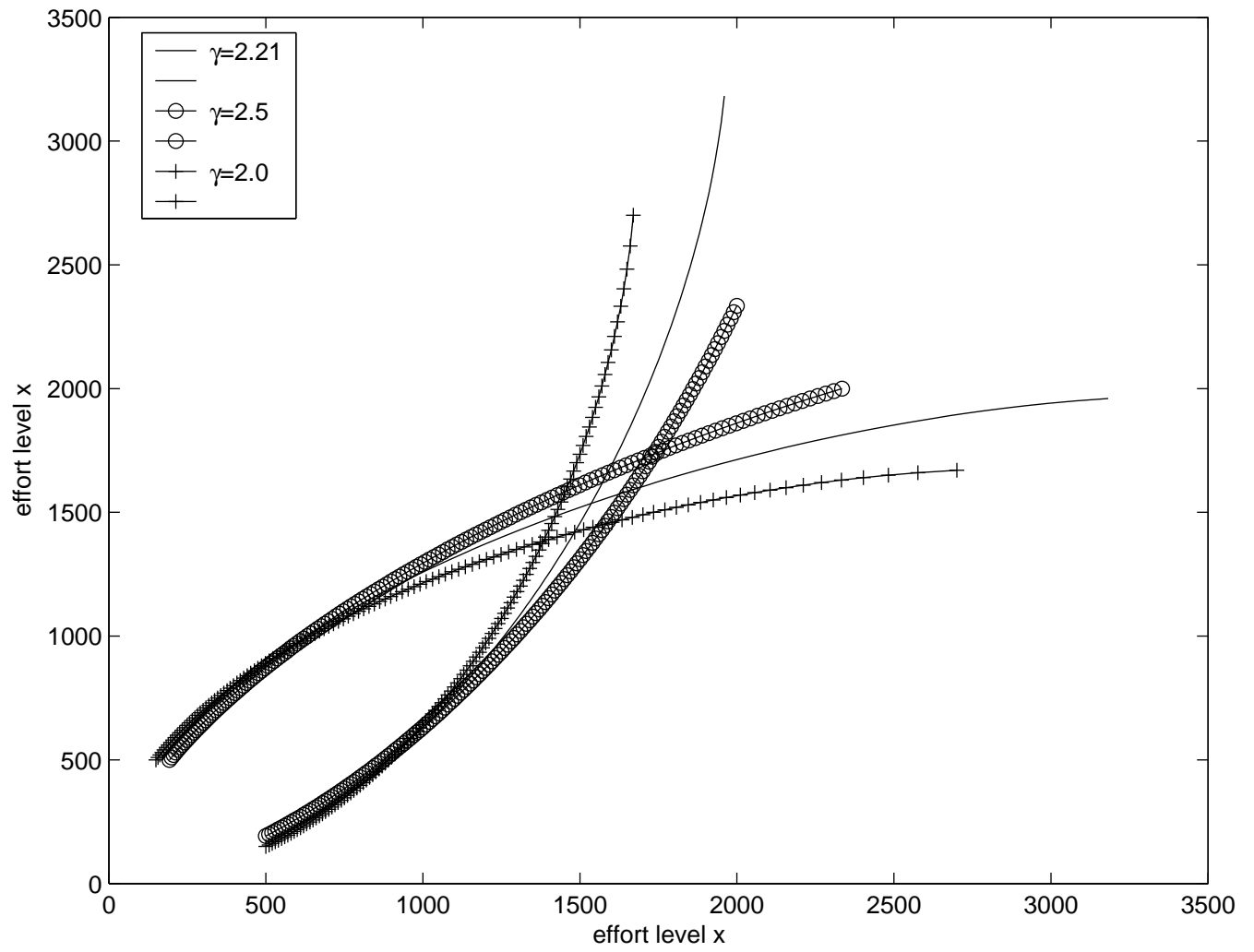


Figure 2: Effects of changes in γ (slope parameter for $h(\cdot)$ function)

Table 1: Average wage and tournament parameters in major retail chains

Store name	# mkt/yr obs.	Dist mgr ($S = 0$)	Store mgr ($S = 1$)	Asst store mgr ($S = 2$)	Sales staff ($S = 3$)
A (Clothing)	13	31978.39	19171.11 1/4.92	13875.25 1.32/3.00	8645.20 3.00/3.48
B (Athletic footwear)	8	37327.43	18051.92 1/16.5	11334.86 1.07/2.19	8016.21 2.19/2.78
C (Footwear)	9	25302.87	14628.8 1/11.56	8048.46 1.11/2.39	6069.92 2.39/3.18
D (High-end specialty)	12	45563.5	24468.35 1/3.5	16019.71 1.5/2.46	9665.78 2.46/3.40
E (Clothing)	38	35767.75	20855.25 1/14.71	14304 1.12/2.47	11025.61 2.47/3.90
F (High-end specialty)	12	41695.05	25456.24 1/7.08	18018.76 1.24/2.10	9943.56 2.10/6.04
G (Children)	13	40410.57	25120.34 1/4.77	18464.37 1.32/4.12	11328.93 4.12/5.89
H (Athletic footwear)	21	29811.7	17989.24 1/8.75	10203.93 1.18/3.71	8403.99 3.71/4.53
I (Clothing)	42	38017.61	21691.99 1/7.33	16480.2 1.19/3.53	10361.4 3.53/6.46
J (Eyewear)	14	45495.18	25910.98 1/9.55	20066.61 1.15/2.11	9382.12 2.11/3.46
K (Clothing)	21	43205.93	24275.2 1/6.21	18040.99 1.27/4.35	12208.09 4.35/7.07
L (Household items)	17	47210.02	32884.84 1/2.94	23046.81 1.79/5.66	10885.75 5.66/13.07
M (Footwear)	45	42655.28	20606.52 1/9.15	10298.17 1.32/3.15	8651.43 3.15/5.20
N (Books)	20	32946.63	19329.51 1/5.5	12181.77 1.35/2.69	8871.47 2.69/4.35

Top entry in each cell is annual salary (in 1986 dollars). Second entry in each cell gives g_s/f_s , the ratio of “winners” from each subgroup to the size of each subgroup. For Sales Staff, salary is calculated as hourly wage*40 hours*50 weeks.

Table 2: Wages, Employment, and Effort Levels for Retailer I, by geographic location (using 1998 data)

Location	Position	Annual salary ^a	# positions	Prob. prom.	Effort Level ^b
Boston	District Manager	35460	10		
	Store Manager	20614	70	0.13	1623.8
	Asst Store Manager	14037	187	0.30	1250.2
	Sales Staff	10098	19	0.93	76.885
Chicago	District Manager	34908	7		
	Store Manager	20184	68	0.09	1246
	Asst Store Manager	13988	109	0.41	1384.6
	Sales Staff	9472	34	0.84	277.72
Dallas	District Manager	37117	9		
	Store Manager	20184	68	0.12	1747.8
	Asst Store Manager	14380	146	0.35	1204.8
	Sales Staff	9755	34	0.87	192.69
Washington DC	District Manager	37117	8		
	Store Manager	19571	68	0.11	1652.5
	Asst Store Manager	13436	144	0.35	1280.5
	Sales Staff	9411	31	0.87	164.97
White Plains	District Manager	37914	6		
	Store Manager	20982	55	0.10	1501.7
	Asst Store Manager	15067	124	0.33	1211.1
	Sales Staff	11644	46	0.80	199.9
Los Angeles	District Manager	37607	6		
	Store Manager	22761	48	0.11	1466.3
	Asst Store Manager	14442	93	0.37	1756.4
	Sales Staff	10282	43	0.77	316.0
Long Island	District Manager	36994	8		
	Store Manager	22270	55	0.13	1632.3
	Asst Store Manager	15939	141	0.31	1224.1
	Sales Staff	10258	69	0.75	330.2
Manhattan	District Manager	43190	8		
	Store Manager	25153	42	0.16	2424.1
	Asst Store Manager	16748	125	0.29	1520.0
	Sales Staff	10908	146	0.55	407.9
Memphis	District Manager	36319	9		
	Store Manager	20368	65	0.12	1704.0
	Asst Store Manager	14074	139	0.35	1300.5
	Sales Staff	9754	56	0.79	266.0
Miami	District Manager	35092	6		
	Store Manager	20368	57	0.10	1268.7
	Asst Store Manager	13607	104	0.38	1465.4
	Sales Staff	10380	20	0.89	142.2
Minneapolis	District Manager	36135	7		
	Store Manager	20429	68	0.09	1329.0
	Asst Store Manager	15202	138	0.35	1109.4
	Sales Staff	10785	57	0.79	276.3
Northern NJ	District Manager	34356	8		
	Store Manager	22025	41	0.16	1684.6
	Asst Store Manager	15374	123	0.28	1189.7
	Sales Staff	10798	50	0.77	223.6
Seattle	District Manager	32393	5		
	Store Manager	20798	36	0.12	1241.6
	Asst Store Manager	14000	104	0.28	1241.2
	Sales Staff	9497	26	0.85	167.5
San Francisco	District Manager	37178	7		
	Store Manager	22515	51	0.12	1556.0
	Asst Store Manager	14675	103	0.36	1632.2
	Sales Staff	11472	66	0.71	292.1

^aDeflated by the US city average CPI, with 1982-84=100. For Sales Staff, salary is calculated as hourly wage*40 hours*50 weeks.

^bObtained by solving Eq. (13).

Assumption: $L_1 = \#(\text{district managers})$, $L_2 = L_3 = \#(\text{store managers})$

Table 4: Average Effort Levels for Retail Chains: First Approach

RETAIL CHAIN	# (Yr-Location) Obs. ^a	Avg x_1	Avg x_2	Avg x_3
A	13	2235.45	1133.35	255.35
B	8	1116.41	1571.09	368.31
C	9	942.82	1480.86	242.50
D	12	4401.48	1596.88	884.16
E	37	1045.65	1463.51	397.77
F	12	2172.51	1538.92	1133.52
G	13	2628.61	1255.02	389.59
H	21	1356.62	1486.62	105.04
I	42	2042.09	1061.45	419.43
J	14	2182.13	1328.67	1405.78
K	21	2886.86	1183.08	418.15
L	17	3224.23	1491.72	609.51
M	45	2705.18	1991.41	725.34
N	20	2315.18	1447.31	462.08

^aWe only included those (Yr-Location) observations for which the calculated effort levels of x_1, x_2, x_3 satisfy the two conditions given in Section xx of the main text.

Table 5: Parameter estimates: Second approach

RETAIL CHAIN	#obs	x_1 : SM effort (std er)	x_2 : ASM effort (std er)	x_3 : Sales effort (std er)	γ^a (std er)	J -statistic ^b (p-value)
A	32	4362.4 (249.27)	832.91 (181.15)	487.94 (78.085)	1.9381 (0.16768)	0.1544 (0.9846)
B	32	2673.1 (680.37)	2262.7 (200.33)	583.59 (54.814)	2.2254 (0.18535)	0.4129 (0.9376)
C	36	2255.5 (120.56)	2183.7 (178.14)	362.74 (36.304)	2.4918 (0.15408)	0.2756 (0.9646)
D	48	3758 (735.92)	2083.7 (870.01)	1493 (88.547)	1.5714 (0.27743)	0.8159 (0.8457)
E	152	2557.5 (1222.1)	1025.1 (506.87)	778.45 (67.762)	2.4557 (0.43442)	0.0445 (0.9975)
F	48	2977.7 (177.49)	2812.9 (416.65)	1600.3 (217.26)	1.851 (0.31762)	0.0609 (0.9961)
G	52	3321.4 (176.44)	2465.8 (618.47)	570.87 (81.442)	1.9835 (0.072259)	0.4706 (0.9253)
H	84	5613 (494.92)	1302 (89.528)	96.461 (33.492)	3.8583 (0.44187)	0.1691 (0.9824)
I	168	4253 (384.81)	1546.6 (377.7)	601.68 (89.842)	2.2125 (0.20805)	0.1439 (0.9861)
J	56	1912.7 (776.83)	3341.8 (2135.6)	1551 (1174.5)	1.3439 (0.22078)	0.1731 (0.9818)
K	84	3779 (944.32)	1930.5 (355.4)	435.37 (33.458)	1.6322 (0.18793)	0.0151 (0.9995)
L	68	4140.3 (1491.5)	3507.3 (1830.9)	1030.9 (326.21)	2.4728 (0.76049)	0.1744 (0.9816)
M	180	3316.1 (2045.7)	1412.9 (3086.2)	1661.4 (1250.4)	2.3372 (1.5903)	0.2271 (0.9731)
N	64	3709.8 (3699.1)	2122 (964.23)	510.78 (116.35)	1.8669 (0.93393)	0.4765 (0.9240)

^aExponent on $h(\cdot)$.^basymptotically distributed $\chi^2(3)$ under null that the moment conditions in Eq. (19) hold.

Table 6: Retailer characteristics

Definitions:*DCLOTH*: =1 if clothing retailer*DFOOT*: =1 if footwear retailer*DHOUSE*: =1 if housewares retailer*DLOC*: =0 if stores mostly located in shopping centers; =1 if in shopping malls*DSIZE*: =0 if stores mostly < 20,000 sq. ft.; =1 if \geq 20,000

RETAIL CHAIN	<i>DCLOTH</i>	<i>DFOOT</i>	<i>DHOUSE</i>	<i>DLOC</i>	<i>DSIZE</i>
A	1	0	0	1	1
B	0	1	0	1	1
C	0	1	0	1	1
D	0	0	1	1	1
E	1	0	0	1	1
F	0	0	1	1	1
G	0	0	0	1	0
H	0	1	0	0	0
I	1	0	0	1	0
J	0	0	0	1	1
K	1	0	0	1	0
L	0	0	1	0	0
M	0	1	0	0	1
N	0	0	0	1	1

Table 7: Regressions of estimated parameters on retailer characteristics

REGRESSORS:	DEPENDENT VAR:	
	γ	$x/1000$ (effort lvl)
<i>Retailer chars:</i>		
<i>DLOC</i>	-0.519 (0.404) ^a	-0.342 (0.398)
<i>DSIZE</i>	-0.328 (0.297)	-0.201 (0.293)
<i>DCLOTH</i>	0.274 (0.345)	-0.318 (0.340)
<i>DFOOT</i>	0.765 (0.404)	-0.345 (0.398)
<i>DHOUSE</i>	0.061 (0.389)	0.319 (0.383)
<i>Hierarchy level:</i>		
Sales		-2.633 (0.288)
ASM		-1.414 (0.289)
Constant	2.469 (0.432)	3.993 (0.458)
<i>N</i>	14	42
<i>R</i> ²	0.665	0.728

^aStandard errors in parentheses. In calculating standard errors, we do not yet take into account fact that dependent variable is estimated.

Table 8: Average percentage of wage differentials accounted for by effort differentials

RETAIL CHAIN	First approach		Second approach ^a	
	$\% \left(\frac{c_1 - c_2}{w_1 - w_2} \right)$	$\% \left(\frac{c_2 - c_3}{w_2 - w_3} \right)$	$\% \left(\frac{c_1 - c_2}{w_1 - w_2} \right)$ (stder)	$\% \left(\frac{c_2 - c_3}{w_2 - w_3} \right)$ (stder)
A	32.57	19.47	65.034 (6.3901)	8.2522 (5.1432)
B	-6.54	37.71	6.3808 (10.079)	52.468 (7.2019)
C	-8.16	78.41	1.1093 (3.7581)	114.53 (12.934)
D	38	12.32	21.441 (17.993)	9.6921 (14.291)
E	-3.56	20.09	28.189 (19.02)	2.6534 (5.762)
F	10.01	4.73	2.3581 (7.0244)	15.306 (7.8527)
G	24.45	12.88	13.628 (11.384)	28.347 (10.158)
H	-2.07	91.43	57.004 (7.3568)	78.045 (7.03)
I	47.95	16.81	78.048 (20.811)	18.824 (9.1563)
J	15.39	-0.15	-26.785 (35.647)	17.531 (16.403)
K	39.52	14.11	35.74 (20.393)	27.166 (6.6007)
L	19.43	7.66	6.8141 (35.144)	21.698 (18.79)
M	7.36	85.55	19.064 (26.851)	-17.171 (157.22)
N	13.18	33.91	23.327 (45.913)	55.884 (32.101)

^aCorresponding to GMM estimates from Table 5.

Table 9: Observed vs. first-best total wage bill implied by estimates

RETAIL CHAIN	First approach			Second approach ^a		
	Observed wage bill (\$mills)	First-best wage bill (\$mills)	% Diff (Obs-FB)/Obs	Observed ^b wage bill (\$mills)	First-best wage bill (\$mills) (stder)	% Diff (Obs-FB)/Obs
A	2.85	1.86	0.347	1.808	1.301 (0.018)	0.28
B	4.35	2.93	0.326	4.356	3.160 (0.104)	0.27
C	3.35	2.24	0.331	3.347	2.498 (0.337)	0.25
D	1.35	0.87	0.356	1.349	0.836 (0.239)	0.38
E	30.02	23.48	0.218	30.344	23.793 (0.934)	0.22
F	5.29	3.78	0.285	5.286	3.872 (0.055)	0.27
G	5.89	4.03	0.316	5.893	4.239 (0.109)	0.28
H	19.32	15.54	0.196	19.318	17.222 (0.197)	0.11
I	178.43	131.85	0.261	178.431	137.952 (1.785)	0.23
J	9.30	5.05	0.457	9.296	5.344 (0.254)	0.43
K	14.98	11.30	0.246	14.977	11.815 (0.192)	0.21
L	6.94	4.82	0.306	6.939	5.044 (0.224)	0.27
M	58.29	42.32	0.274	58.289	41.084 (3.123)	0.30
N	4.37	3.30	0.245	3.343	2.714 (0.324)	0.19

^aCorresponding to GMM estimates from Table 5.

^bNote that figures in first and third columns may not coincide due to (i) rounding errors; and (ii) for some (firm/geographic locations/year) observations, we were not able to obtain convergent estimates for the second approach.

Wages bills only for store managers, assistant store managers, and sales staff.

Table 10: Parameter estimates: incorporating turnover rates

RETAIL CHAIN	#obs	x_1 (std er)	x_2 (std er)	x_3 (std er)	γ^a (std er)	J -statistic ^b (p-value)
A	32	3322.3 (506.02)	3588.8 (2106.6)	207.81 (504.02)	1.107 (0.238)	1.853 (0.603)
B	32	1822.8 (649.07)	3733.1 (5201.8)	1861.4 (6434.5)	1.822 (0.276)	1.887 (0.596)
G	24 ^c	3614.6 (772.51)	2448.2 (10047)	691.47 (1387.5)	0.918 (0.056)	2.018 (0.569)
M	60	3057.6 (1497.4)	5681.9 (4339.5)	2549.3 (23681)	1.128 (0.187)	1.730 (0.630)

^aExponent on $h(\cdot)$, link function specification.

^basymptotically distributed $\chi^2(2)$ under null that the moment conditions in Eq. (19) hold.

^cFewer observations are available for this retail chain (as compared to the number of observations used for the results in Table 5) because turnover rates were not reported for some years. Same applies for Retail Chain M.

Table 11: Average percentage of wage differentials accounted for by effort differentials: incorporating turnover rates

RETAIL CHAIN	Second approach ^a	
	$\% \left(\frac{c_1 - c_2}{w_1 - w_2} \right)$ (stder)	$\% \left(\frac{c_2 - c_3}{w_2 - w_3} \right)$ (stder)
A	-4.9105 (36.219)	80.879 (61.988)
B	-29.701 (82.244)	58.486 (143.26)
G	20.257 (183.45)	33.765 (177.58)
M	-25.35 (53.648)	230.93 (2057.6)

^aCorresponding to GMM estimates from Table 10.

Table 12: Observed vs. first-best total wage bill implied by estimates: incorporating turnover rates

RETAIL CHAIN	Second approach ^a		
	Observed ^b wage bill (\$mills)	First-best wage bill (\$mills) (stder)	% Diff (Obs-FB)/Obs
A	1.808	1.455 (0.189)	0.19
B	4.356	2.923 (1.292)	0.33
G	1.478	1.190 (0.301)	0.19
M	37.245	27.203 (43.21)	0.27

^aCorresponding to GMM estimates from Table 10.

^bNote that figures in first and third columns may not coincide due to (i) rounding errors; and (ii) for some (firm/geographic locations/year) observations, we were not able to obtain convergent estimates for the second approach.

Wages bills only for store managers, assistant store managers, and sales staff.