Interpreting Intra-firm Wage Differentials using Tournament Models*

Matthew Shum
Dept. of Economics
Johns Hopkins University

Draft: November 9, 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Please do not quote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.

---

[*] 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 Please do not quote.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, Quang Vuong, and seminar participants at Northwestern and SITE for useful suggestions. Please contact me at Dept. of Economics, Johns Hopkins University, Baltimore, MD 21218; e-mail `mshum@jhu.edu`.

Further up the hierarchy, district managers make on average over \$38,000 (in 1986 dollars).[1] Why do these differentials arise? One explanation is that employees at higher levels of a firm are paid more, because they work harder, or are more productive. Alternatively, the tournament literature proposes that 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)). One common feature of these papers is their focus on industries for which productivity measures at the individual worker level are observable.

However, for many industries, these measures are difficult to obtain (or unavailable) in practice. In this paper, we consider the structural estimation of tournament models when 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 values for the structural elements of the tournament model.

Most closely related to the present paper are two 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.[2]

We illustrate our methodologies by estimating tournament models 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

---

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

[2] Relatedly, Ferrall and Smith (1999) develop and estimate a structural tournament-like model of sports championship series.

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 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. Using our estimates, we are also able to simulate "first-best" compensation levels (i.e., what employees would be paid if effort were observable and contractible), and compare them with actual compensation levels (which we interpret as information-constrained "second-best" wages).

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.[3] As Rosen (1986) points out, intra-firm wage

---

[3] Another 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

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 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.[4] 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 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. $W_1, \ldots, W_{S+1}$ denote the payoffs

---

(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.

[4] Tournament (or compensation) design is typically motivated by the presence of *asymmetric* information between workers and firms. In a symmetric information situation, firms can 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. See Teulings (1995) for structural estimates from a perfect information matching (assignment) model of the labor market.

(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.

Let $n_0, \ldots, n_S$ 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 $s = 0, \ldots, S$.

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.[5]

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.[6] 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

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

---

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

[6] It 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.

In the above, $h(\cdot)$ is a link function translating individual effort levels into the probability of advancement.[7] The $P_s(\cdots)$ function captures, in reduced-form, the procedure whereby 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).[8]

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^*:

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

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, in equilibrium, 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 requirement because, in any symmetric 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.[9]

---

[7] The 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. Th logit probability function can, in turn, 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 $f_s \neq 2$.

[8] However, 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 $P_s(\cdots)$ is explicitly derived from the information structure of the game.

[9] Additionally, 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.

2.2 Equilibrium

The equilibrium sequence of effort levels $\{x_s^* : s = 1, \dots, S\}$ is determined by a dynamic optimization problem.[10] 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 equation[11]

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

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:

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

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

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

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

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

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.)

---

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

[11] Implicitly, we are assuming a discount rate equal to 1. Also, we assume workers are risk-neutral, so that in this model there is no insurance aspect to contracting.

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.[12] 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 the top graph 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. 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 $x_1^*, \ldots, x_S^*$ 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.

---

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

3 Empirical strategy: two approaches

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

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

where

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

In matrix notation, this is

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

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

$$V_s=\beta_s V_{s-1}+(1-\beta_s)\,W_{s+1},\;\; s=1,\ldots,S \tag{10}$$

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

$$\begin{aligned} V_1 &= \beta_1 W_1+(1-\beta_1)W_2 \\ V_2 &= \beta_2\beta_1 W_1+\beta_2(1-\beta_1)W_2+(1-\beta_2)W_3 \\ &\vdots \\ V_S &= \beta_1\ldots\beta_S W_1+(1-\beta_1)\beta_2\ldots\beta_S W_2+(1-\beta_2)\beta_3\ldots\beta_S W_3+\cdots+(1-\beta_S)W_{S+1} \end{aligned}$$

or

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

Substituting (11) into (9), we obtain

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

The system of equations (12) gives us, for each firm, S equations but $2S$ unknowns $c(x_1^*), \ldots, c(x_S^*)$, $\mu(x_1^*), \ldots, \mu(x_S^*)$. 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.

In both approaches we assume:

$$\begin{aligned} h(x) &= x^\gamma \\ c(x) &= x. \end{aligned} \tag{13}$$

That is, we assume a power (constant-elasticity) specification for the h function, and normalize the cost of effort function to be the identity function.[13] Subsequently, a total of $S+1$ parameters— $\theta \equiv (x_1^*, \ldots, x_S^*, \gamma)$ are to be estimated. 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, which may imply, in turn, that a better monitoring technology is in place.

With these assumptions, we can simplify $\mu(x_s^*) = \gamma$ for $s = 1, \ldots, S$, and the system of first-order conditions (19) reduces to

---

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

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

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.

$$\begin{pmatrix} \frac{\gamma}{f_1}(1-p_1^*) & 0 & \cdots & 0 \\ 0 & \frac{\gamma}{f_2}(1-p_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\gamma}{f_S}(1-p_S^*) \end{pmatrix} *$$

$$\left[\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \beta_1 & (1-\beta_1) & \cdots & 0 \\ \beta_1\beta_2 & \beta_2(1-\beta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1\cdots\beta_{S-1} & (1-\beta_1)\beta_2\cdots\beta_{S-1} & \cdots & (1-\beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_S \end{pmatrix} - \begin{pmatrix} W_2 \\ W_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] = \begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_S^* \end{pmatrix} \tag{14}$$

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 $h(\cdot)$. The bottom figure in Fig. 1 illustrates how changes in γ affects equilibrium incentives to exert effort. The solid lines are a pair of best-response curves corresponding to $\gamma = 1.5$. When γ is increased to 2.0, 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 1.0 (illustrated by the best-response curves marked with crosses), the equilibrium effort levels decrease by around 15%.

3.1 First approach

In the first approach, we address the under-identification issues by adding additional equations regarding the supply-side of the model.[14] Specifically, we assume that the employment at the S levels of the firms are chosen to maximize profits at the (chain, year, and geographic location)-level. In specifying the supply-side, we assume that firms set $n_1, \dots, n_S$ optimality, but take the wages $W_1, \dots, W_{S+1}$, as well as the functional form of the advancement

---

[14] Here we depart from Rosen's (1986) model, which does not have a supply-side.

probability function $P_s(\cdots)$ as given. (This is similar to the approach taken in Ferrall (1996).)

We assume that firms choose $n_1, \dots, n_S$ to maximize the following profit objective:

$$\max_{n_1,\dots,n_S} \theta \left(\sum_{s=1}^{S} \left(n_s x_s^*(\vec{n}, \vec{W}) \right)^{\rho} \right)^{1/\rho} - \sum_{s=1}^{S} n_s W_{s+1} \tag{15}$$

where $\vec{n} = (n_1, \dots, n_S)$ and $\vec{W} = (W_1, \dots, W_{S+1})$. In the above, we assume that firm revenues are generated by a CES function, in the aggregate effort levels exerted in the different levels of the company. The substitution parameter ρ should lie in $[-\infty, 1]$ for the production function to be concave, and the multiplicative parameter θ is included as a scaling factor from effort to revenue units. The first-order conditions for this problem imply:

$$\frac{W_{s+1}}{W_{S+1}} = \frac{(n_s x_s^*)^{\rho-1} x_s^* + \sum_{t=1}^{S} (n_t x_t^*)^{\rho-1} n_t \frac{\partial x_t^*}{\partial n_s}}{\left(n_S x_S^*\right)^{\rho-1} x_S^* + \sum_{t=1}^{S} (n_t x_t^*)^{\rho-1} n_t \frac{\partial x_t^*}{\partial n_S}}, \quad s = 1, \dots, S-1 \tag{16}$$

which yield $S-1$ equations, with the extra parameter ρ. (By dividing two first-order conditions, we eliminate the scaling parameter θ, which we are not interested in estimating.)

To simplify, note from examination of the system (14) that x_s^*, for $s = 2, \dots, S$ is only dependent on $n_1, \dots, n_s$. Hence $\frac{\partial x_t^*}{\partial n_s} = 0$ for $s \geq t$, and the above can be simplified to

$$\frac{W_{s+1}}{W_{S+1}} = \frac{(n_s x_s^*)^{\rho-1} x_s^* + \sum_{t=s}^{S} (n_t x_t^*)^{\rho-1} n_t \frac{\partial x_t^*}{\partial n_s}}{\left(n_S x_S^*\right)^{\rho-1} x_S^*}, \quad s = 1, \dots, S-1. \tag{17}$$

For our empirical work, $S = 3$, so that the supply-side adds two equations. These two equations, combined with the S equations given in (14), would allow us to estimate the equilibrium effort levels $x_1^*, \dots, x_S^*$, the power parameter γ of the $h(\cdot)$ function, as well as the new ρ parameter of the CES production function.

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 potentially restrictive assumptions regarding the supply-side. In the second approach, we dispense with the supply-side assumptions, 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.

However, once we assume that $x_1, \ldots, X_S$ and γ remain unchanged over time and geographic locations for each retail chain, we must introduce additional randomness into the model to accommodate the variation in employment and wages levels across time and locations which we observe in the data. Therefore, we do this by allowing the wages $W_0, \ldots, W_S$ to be observed with error.[15]

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}:

$$W_{ismt} = W^*_{ismt} + \epsilon_{ismt}, \;\; s = 1, \ldots, S \tag{18}$$

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 $x_1^*, \ldots, x_S^*$, 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 (18) to obtain

$$\begin{pmatrix} \frac{h'(x_1^*)/h(x_1^*)}{f_1}(1-p_1^*) & 0 & \cdots & 0 \\ 0 & \frac{h'(x_s^*)/h(x_s^*)}{f_2}(1-p_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{h'(x_S^*)/h(x_S^*)}{f_S}(1-p_S^*) \end{pmatrix} *$$
$$\left[\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \beta_1 & (1-\beta_1) & \cdots & 0 \\ \beta_1\beta_2 & \beta_2(1-\beta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1 \cdots \beta_{S-1} & (1-\beta_1)\beta_2 \cdots \beta_{S-1} & \cdots & (1-\beta_{S-1}) \end{pmatrix} \begin{pmatrix} W_1 - \epsilon_1 \\ W_2 - \epsilon_2 \\ \vdots \\ W_S - \epsilon_S \end{pmatrix} - \begin{pmatrix} W_2 - \epsilon_2 \\ W_3 - \epsilon_3 \\ \vdots \\ W_{S+1} \end{pmatrix} \right] = \begin{pmatrix} c'(x_1^*) \\ c'(x_2^*) \\ \vdots \\ c'(x_S^*) \end{pmatrix} \tag{19}$$

---

[15] 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.

or, in shorthand,

$$\begin{aligned} \mathbf{A}\left[\mathbf{B}\left(\vec{W}_{1:S}-\vec{\epsilon}_{1:S}\right)-\left(\vec{W}_{2:S+1}-\vec{\epsilon}_{2:S}|0\right)\right] &= \vec{c}' \Rightarrow \\ \vec{\epsilon} = -\tilde{\mathbf{B}}^{-1}\left[\mathbf{A}^{-1}\vec{c}' - \mathbf{B}\vec{W}_{1:S} + \vec{W}_{2:S+1}\right] & \end{aligned} \tag{20}$$

where $\tilde{\mathbf{B}}$ is the matrix $\mathbf{B}$ minus a $S \times S$ matrix with ones in the $(i, i+1)$, $i = 1, \ldots, S-1$ 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 $\epsilon_2, \ldots, \epsilon_S$ and the S-th element is a zero.) Because the system of equations (19) is only S-dimensional, we cannot accommodate an additional measurement error in the wage W_{S+1}.[16]

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

$$E[\epsilon Z] = 0 \tag{21}$$

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. (20). Let $M \geq S+2$ be the total number of moments conditions employed in estimating θ. We seek the minimizer of the quadratic form

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

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

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

where

$$\begin{aligned} J &= E_0 \frac{\partial \mathbf{m}(\theta_0)}{\partial \theta_0} \\ V &= Var_0\ \mathbf{m}(\theta_0) = E_0 \mathbf{m}(\theta_0)\mathbf{m}(\theta_0)' \end{aligned}$$

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 $(J'\Omega J)^{-1}$.

---

[16] Since 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.

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 $x_1^*, \ldots, x_S^*$ to ensure that they satisfy the second-order condition which corresponds to the first-order condition in Eq. (4):

$$\hat{c}'(\hat{x}_s^*) * \left[\frac{\hat{h}''(\hat{x}_s^*)}{\hat{h}'(\hat{x}_s^*)} - 2\frac{\hat{h}'(\hat{x}_s^*)}{f_s * \hat{h}(\hat{x}_s^*)} \right] - \hat{c}''(\hat{x}_s^*) < 0 \tag{22}$$

for $s = 1, \ldots, S$, and the hats ($\hat{}$) 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 typically 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. According to the tournament model, we interpret the observed number of workers at stages $s = 1, \ldots, S$ as the number of "losers" (n_s) at that stage. The number of district managers, n_0, is interpreted as the number of "winners" in stage $s = 1$ of the tournament game.[17]

At the sales staff, assistant store manager, and store manager hierarchical levels, 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 full-time 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.[18]

Twelve of the fourteen chains studied in this analysis report that the compensation of

---

[17] We do not model the changes across years in employment at the (chain, geographic location)-level. This would require extra assumption and modeling of the flow of workers in and out of the firm, which is beyond the scope of this paper (especially given that we have no data on the job tenure of the workers at each hierarchical level).

[18] Note 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.

their sales-staff are "straight salary" meaning that less than 25% of their compensation is commission-based. The three exceptions are retailer B in 1999[19], retailer C in 1998[20], and retailer F in 1999[21]. These stores reported that the sales staff were compensated in a fashion such that over 25% of the compensation arose from commissions.

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 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.]

---

[19] the only year we have data for this retailer

[20] In 1997, this retailer reported that its sales-staff compensation was "straight salary", and no data was available for this retailer in 1999.

[21] In 1998, this retailer reported that sales staff were compensated via "straight salary", while no data was available in 1997.

5 Empirical results

5.1 First approach

Using the first approach, we are able to recover the values of γ, ρ, x_1^*, x_2^*, x_3^* for every (year-location) observations of $(W_1, \ldots, W_{S+1}, n_0, \ldots, n_S)$, for each retail chain. Table 2 presents the average and standard deviation of the recovered values for each retailer, where the average is taken across all (year, geographic region) pairs observed in the dataset for that particular chain.[22]

[Table 2 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. However, for several retailers, the average effort level drops at higher hierarchical levels: for example, for retailers B and C (which, incidentally, are both footwear retailers) the average effort level at the assistant store manager stage exceeds the average effort level at the store manager stage. 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.

[Table 3 about here.]

Next, we see that the average $\gamma > 1$ in all the retail chains, excepting chain J (which is the sole eyewear retailer in our dataset). The implies that for most chains, the $h(\cdot)$ function, which measures the sensitivity of the promotion probability to effort levels, is convex on average. Finally, for a large number of the retailers, the average ρ is quite negative, implying a low elasticity of substitution between the various types of workers which we consider.

Because we are able to recover effort levels, as well as γ and ρ, at the (retailer, year, geographic region)-level, we next examine the variation in these quantities. To begin with,

---

[22] As discussed above, for each set of estimates, we checked that the second-order conditions in Eqs. (22). We also dropped observations for which the nonlinear equation solver did not converge, as well as observations where the recovered $\rho > 1$, implying that the production function is not concave. Among these criteria, the non-convergence of the nonlinear solver accounted for the most eliminations – 56% of the observations.

a simple correlation table (not reported here for brevity) for the recovered quantities shows that all the effort levels x_1, x_2, and x_3 are positively correlated with γ (and with each other), but that there is little discernible correlation between the ρ's and the other quantities.

Next, we ran some regressions to see whether store characteristics could explain the variation in these quantities. For each retailer, we created dummy variables, for line of business and also location and size characteristics (which are defined and summarized in Table 3). Subsequently, we performed regressions of the estimated γ's, ρ's and effort levels on these dummy variables.

The regression results are reported in Table 4. For γ, the parameter in the $h(\cdot)$ function linking effort to the probability of winning, we see that most of the coefficients are marginally significant (at around a 10-20% significance level). γ tends to be smaller in shopping malls (vs. shopping centers) and in larger stores. Curiously, the coefficients are negative across the line-of-business dummies $DCLOTH$, $DFOOT$, and $DHOUSE$ (relatively to the omitted category, which consists of the children's retailer G and the eyewear retailer J). Interestingly, we find that γ is significantly higher in stores which pay sales staff using commissions. To the extent that a larger γ is due to less noisy observations of effort (which in turn could be due to higher monitoring intensity), one explanation of this relationship may be that there exist complementarities between monitoring and offering commissions in the objective function of the retailer.

In the regressions where ρ, the CES production parameter, is the dependent variable, none of the coefficients (except the positive coefficient on $DHOUSE$) are precisely estimated. In the three effort regressions, the most striking finding is that $DCOMMIS$ enters positively (and significantly) in the regression where x_3, the sales staff effort, is the dependent variable, suggesting that offering commissions to sales staff increase their effort.[23] Since we do not impose this in recovering the effort levels, the finding of the expected positive relationship between $DCOMMIS$ and x_3 offers additional support for our model, and our interpretation of the x's as effort levels.

[Table 4 about here.]

---

[23] We do not include $DCOMMIS$ as a regressor in the x_1 and x_2 regressions, because the survey only asks retailers to report whether commissions were used for sales staff.

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.[24] 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 5 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 conditions (in Eq. (22)) 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 retailers[25], ranging from a low value of 1.34 for retailer J, to 3.86 for retailer H. Furthermore, for a majority of the choices, the estimated γ using the second approach is higher than the average recovered using the first approach. Correspondingly, it is not surprising to find that, generally, the effort levels estimated using the second approach are higher than those reported earlier because, as we discussed above, equilibrium effort levels increase when γ increase.

Furthermore, a 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

---

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

[25] Given 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.

where sales staff are required to have the most training, as their duties include fitting and ordering eyewear for customers.

As with the first approach results, we also ran regressions of the estimated γ's and effort levels on retailer characteristics, as reported in Table 6. In the γ regression, none of the coefficients are statistically distinguishable from zero, which is not surprising given the small number of observations. Because only one set of parameters are estimated for each retail chain using the second approach, we pool the effort levels across all hierarchical levels in running the regressions with effort as the dependent variable, which are reported in the third column of Table 6. As expected, the hierarchical dummies enter significantly, reflected the fact that higher effort is exerted at higher levels of the hierarchy. However, none of the other covariates enter significantly.

[Table 6 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 7 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. We focus on the first approach results (reported in the leftmost columns of Table 7) in the following discussion.

[Table 7 about here.]

All the estimated δ's are less than 100, and typically less than 50. 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 (retailers B, C, and E between the assistant store manager and store manager stages), 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.[26]

These wage bill results are given in Table 8. 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 8 for, respectively, the first and second approach estimates) are not unreasonably high, hovering between 20-25% for most retailers, and ranging from 19.5% for retailer M, up to 46.7 for retailer J (using the first 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.[27]

[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 several ways to recover 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 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

---

[26] We 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).

[27] Ideally, 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.

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.

Finally, our dataset also contains a wealth of data on compensation-related contractual features within the retail chains, especially at higher levels of the firm. While these have not been exploited and used in this paper, there is a sizable theoretical literature regarding optimal contracting within a firm. For instance, Holmstrom and Milgrom (1994) provide a theoretical framework to analyze statistical associations between a firm's endogenous design variables. It will be interesting to use this framework to interpret the contractual features observed for the firms in our dataset, and draw inferences regarding the shapes of firms' objective functions.

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.

HOLMSTROM, B., AND P. MILGROM (1994): “The Firm as an Incentive System,” *American Economic Review*, 84, 972–991.

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.

TEULINGS, C. (1995): “The Wage Distribution in a Model of the Assignment of Skills to Jobs,” *Journal of Political Economy*, 103, 280–315.

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 $\rho_0, \ldots, \rho_S$, 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).[28] 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, $s = 0, \ldots, S$ denote the CDF of outside wages for employees in stage s. We can interpret the observed turnover rates as

$$\rho_s = 1 - F_s(W_{s+1}), \;\; s = 0, \ldots, S. \tag{23}$$

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

$$V_s = \max_x \left\{ p_s(x; x_s^*) V_{s-1} + (1 - p_s(x; x_s^*)) \, E_{R_s} \max \left(W_{s+1}, R_s \right) \right\} \tag{24}$$

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

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

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

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

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

---

[28] Obviously, 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:

$$\begin{aligned} F_s(W_{s+1}) &= \frac{W_{s+1}}{\bar{\theta}_s} = 1 - \rho_s \\ \Rightarrow \ \bar{\theta}_s &= \frac{W_{s+1}}{1 - \rho_s}, \ s = 0, \dots, S \end{aligned} \tag{25}$$

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

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

Hence, after plugging in these items into Eq. (24), we can estimate as before.

For the first approach, given observations of $\rho_0, \dots, \rho_S$, we can back out $\bar{\theta}_0, \dots, \bar{\theta}_S$ using Eq. (25). Then we can construct $\tilde{W}_1, \dots, \tilde{W}_{S+1}$ and then estimate $x_1^*, \dots, x_S^*$, as well as γ and ρ, using the system of equations (14), 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:[29]

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

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. (20) in the main text is

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

where

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

---

[29] As 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 $s = 0, \ldots, S$):

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

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

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 9. 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.[30]

However, the qualitative implications of the results are similar to the above results. The results in Tables 10 and 11 are the turnover model analogs to those reported in, respectively, Tables 7 and 8 for the no-turnover model.

[Table 9 about here.]

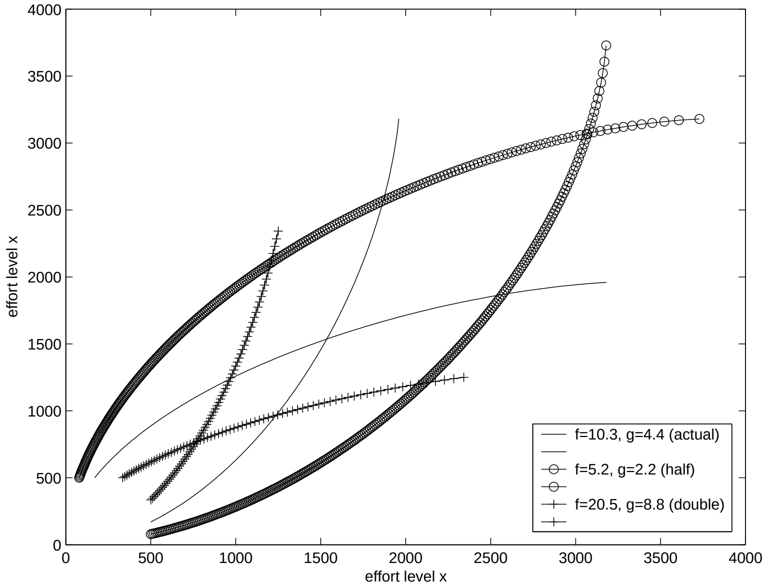
[Table 10 about here.]

[Table 11 about here.]

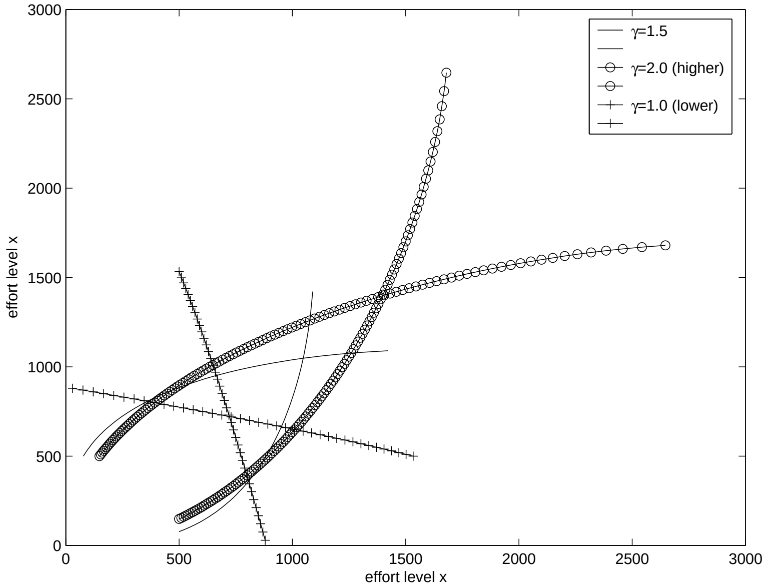
---

[30] However, the second-order conditions in Eq. (22) continue to hold for these estimates.

Figure 1: Comparative statics of tournament model



Effects of changes in toughness of competition



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

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

Store identifier	# 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	13875.25	8645.20
		(3968.13)	(1550.75)	(1835.10)	(894.61)
		1.23	4.69	7.77	2.08
			1/4.92	1.32/3.00	3.00/3.48
B (Athletic footwear)	8	37327.43	18051.92	11334.86	8016.21
		(5421.04)	(1970.93)	(553.25)	(495.38)
		1.00	15.50	17.00	9.00
			1/16.5	1.07/2.19	2.19/2.78
C (Non-athletic Footwear)	9	25302.87	14628.8	8048.46	6069.92
		(2654.97)	(916.31)	(348.82)	(1755.60)
		1.22	13.33	16.22	9.11
			1/11.56	1.11/2.39	2.39/3.18
D (High-end specialty)	12	45563.5	24468.35	16019.71	9665.78
		(4838.46)	(4261.09)	(1415.18)	(999.65)
		1.00	2.50	2.17	1.83
			1/3.5	1.5/2.46	2.46/3.40
E (Clothing)	38	35767.75	20855.25	14304	11025.61
		(6502.10)	(2998.42)	(2310.14)	(1196.92)
		1.05	14.58	18.32	21.87
			1/14.71	1.12/2.47	2.47/3.90
F (High-end specialty)	12	41695.05	25456.24	18018.76	9943.56
		(7138.76)	(3971.62)	(2238.46)	(1358.79)
		1.00	6.08	5.33	18.50
			1/7.08	1.24/2.10	2.10/6.04
G (Children)	13	40410.57	25120.34	18464.37	11328.93
		(7636.94)	(4064.84)	(1916.36)	(1006.76)
		1.23	4.54	12.38	7.92
			1/4.77	1.32/4.12	4.12/5.89
H (Athletic footwear)	21	29811.7	17989.24	10203.93	8403.99
		(3389.65)	(1799.55)	(1013.59)	(748.64)
		2.48	19.76	49.24	9.33
			1/8.75	1.18/3.71	3.71/4.53
I (Clothing)	42	38017.61	21691.99	16480.2	10361.4
		(3214.85)	(1720.96)	(3088.39)	(714.72)
		8.36	50.81	111.07	126.62
			1/7.33	1.19/3.53	3.53/6.46
J (Eyewear)	14	45495.18	25910.98	20066.61	9382.12
		(4900.57)	(2287.53)	(2478.05)	(600.92)
		1.36	11.71	10.79	15.57
			1/9.55	1.15/2.11	2.11/3.46
K (Clothing)	21	43205.93	24275.2	18040.99	12208.09
		(4135.74)	(2255.50)	(1466.40)	(1175.39)
		1.52	7.81	19.10	14.95
			1/6.21	1.27/4.35	4.35/7.07
L (Household items)	17	47210.02	32884.84	23046.81	10885.75
		(5267.12)	(3002.86)	(2175.07)	(2252.73)
		1.00	1.94	7.06	16.41
			1/2.94	1.79/5.66	5.66/13.07
M (Non-athletic Footwear)	45	42655.28	20606.52	10298.17	8651.43
		(4090.03)	(2132.92)	(974.77)	(703.60)
		3.76	32.73	27.31	43.73
			1/9.15	1.32/3.15	3.15/5.20
N (Books)	20	32946.63	19329.51	12181.77	8871.47
		(3571.39)	(2123.77)	(1400.18)	(971.85)
		1.15	4.75	5.50	6.40
			1/5.5	1.35/2.69	2.69/4.35

Top entry in each cell is annual salary (in 1986 dollars). Second entry is standard deviation of salary. Third entry in each cell gives n_s, and fourth entry 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: Average Parameters for Retail Chains: First Approach

RETAIL CHAIN	# (Yr-Location) Obs.[a]	Avg x_1 (stdev)	Avg x_2 (stdev)	Avg x_3 (stdev)	Avg γ (stdev)	Avg ρ (stdev)
A	6	4556.08 (2322.71)	1639.03 (870.62)	398.82 (172.03)	2.03 (1.07)	-1.56 (0.46)
B	8	2309.14 (991.30)	2521.79 (689.53)	527.70 (203.40)	2.07 (0.65)	-1.71 (0.82)
C	5	1388.00 (788.33)	1651.83 (468.95)	209.73 (55.91)	1.15 (0.32)	-0.83 (0.29)
D	4	9063.02 (6284.87)	1208.78 (865.31)	491.68 (381.74)	1.50 (1.00)	-0.11 (1.19)
E	25	1315.08 (905.33)	1967.67 (1443.30)	542.39 (433.93)	1.41 (1.04)	-8.33 (28.35)
F	0[b]	—	—	—	—	—
G	3	1989.53 (1800.92)	1024.25 (1065.05)	259.74 (298.83)	1.04 (1.25)	-2.43 (1.48)
H	6	2020.78 (2084.50)	1454.38 (489.05)	215.72 (91.43)	1.05 (0.40)	-1.09 (0.53)
I	20	2373.51 (836.12)	1516.83 (397.51)	299.66 (134.12)	1.36 (0.47)	-2.45 (0.77)
J	2	280.72 (1147.47)	768.11 (1071.53)	852.58 (1194.34)	0.55 (0.75)	-10.74 (2.57)
K	13	3523.32 (1994.66)	1314.03 (896.51)	372.74 (408.95)	1.32 (1.04)	-3.83 (4.76)
L	2	5533.15 (1646.92)	2993.55 (971.07)	1787.60 (327.81)	2.69 (0.90)	-3.34 (0.74)
M	23	5488.56 (4237.27)	2656.54 (1105.44)	500.64 (230.57)	1.82 (0.75)	-2.18 (1.56)
N	14	5409.00 (2794.65)	2058.66 (720.71)	651.07 (190.33)	2.31 (0.64)	-3.49 (3.82)

---

[a] We only included those (Yr-Location) observations for which the calculated effort levels of x_1, x_2, x_3 satisfy the two conditions given at the end of Section 3.2 of the main text, and the γ and ρ parameters were obtained as zeros of the supply-side first-order-conditions (17).

[b] There were no (year-location) observations which both satisfied the two model optimality conditions, as well as yielded convergent estimates for the nonlinear solver used to solve the supply-side first-order conditions.

Table 3: 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 $\geq 20{,}000$
$DCOMMIS$: =1 if over 25% of sales-staff pay is based on commissions

RETAIL CHAIN	$DCLOTH$	$DFOOT$	$DHOUSE$	$DLOC$	$DSIZE$	$DCOMMIS$
A	1	0	0	1	1	0
B	0	1	0	1	1	1[a]
C	0	1	0	1	1	1[b]
D	0	0	1	1	1	0
E	1	0	0	1	1	0
F	0	0	1	1	1	1[c]
G	0	0	0	1	0	0
H	0	1	0	0	0	0
I	1	0	0	1	0	0
J	0	0	0	1	1	0
K	1	0	0	1	0	0
L	0	0	1	0	0	0
M	0	1	0	0	1	0
N	0	0	0	1	1	0

---

[a] Only for 1999
[b] Only for 1998
[c] Only for 1999

Table 4: Regressions of parameters on retailer characteristics: first approach with supply-side

	DEPENDENT VAR:				
REGRESSORS:	γ	ρ	$x_1/1000$	$x_2/1000$	$x_3/1000$
DLOC	-0.778	1.831	-2.085	-0.639	-0.513
	(0.391)	(5.921)	(0.956)	(0.321)	(0.151)
DSIZE	0.341	-3.591	0.698	0.564	0.182
	(0.177)	(2.686)	(0.612)	(0.206)	(0.069)
DCLOTH	-0.338	-0.055	-2.216	0.255	-0.103
	(0.256)	(3.877)	(0.885)	(0.297)	(0.099)
DFOOT	-0.873	3.460	-2.444	0.354	-0.574
	(0.405)	(6.139)	(1.113)	(0.374)	(0.157)
DHOUSE	-0.488	12.294	2.134	-0.271	0.069
	(0.439)	(6.658)	(1.517)	(0.510)	(0.170)
DCOMMIS	1.097	-3.754	—	—	0.520
	(0.511)	(7.744)	(—)	(—)	(0.198)
year dummies	yes	yes	yes	yes	yes
city dummies	yes	yes	yes	yes	yes
N	131	131	131	131	131
R^2	0.305	0.224	0.316	0.341	0.388

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

---

[a] Exponent on $h(\cdot)$.

[b] asymptotically distributed $\chi^2(3)$ under null that the moment conditions in Eq. (21) hold.

Table 6: Regressions of estimated parameters on retailer characteristics

	DEPENDENT VAR:	
REGRESSORS:	γ	$x/1000$ (effort lvl)
Retailer chars:		
$DLOC$	-0.421	-0.343
	(0.667)[a]	(0.432)
$DSIZE$	-0.341	-0.202
	(0.324)	(0.298)
$DCLOTH$	0.271	-0.318
	(0.368)	(0.346)
$DFOOT$	0.889	-0.346
	(0.776)	(0.448)
$DHOUSE$	0.093	0.319
	(0.448)	(0.393)
$DCOMMIS$	-0.148	
	(0.771)	
Hierarchy level:		
Sales		-2.634
		(0.309)
ASM		-1.414
		(0.293)
$DCOMMIS*$Sales		0.003
		(0.688)
Constant	2.379	3.994
	(0.654)	(0.495)
N	14	42
R^2	0.667	0.728

---

[a] Standard errors in parentheses. In calculating standard errors, we do not yet take into account fact that dependent variable is estimated.

Table 7: Average percentage of wage differentials accounted for by effort differentials

	First approach[a]		Second approach[b]	
RETAIL CHAIN	$\% \left(\frac{c_1-c_2}{w_1-w_2}\right)$ (stdev)	$\% \left(\frac{c_2-c_3}{w_2-w_3}\right)$ (stdev)	$\% \left(\frac{c_1-c_2}{w_1-w_2}\right)$ (stder)	$\% \left(\frac{c_2-c_3}{w_2-w_3}\right)$ (stder)
A	51.25	32.25	65.034	8.2522
	(36.64)	(12.21)	(6.3901)	(5.1432)
B	-2.12	60.86	6.3808	52.468
	(17.05)	(20.66)	(10.079)	(7.2019)
C	-3.90	93.43	1.1093	114.53
	(8.26)	(18.24)	(3.7581)	(12.934)
D	100.51	14.45	21.441	9.6921
	(75.81)	(12.50)	(17.993)	(14.291)
E	-4.79	23.99	28.189	2.6534
	(19.99)	(156.10)	(19.02)	(5.762)
F			2.3581	15.306
			(7.0244)	(7.8527)
G	14.42	15.43	13.628	28.347
	(12.53)	(13.53)	(11.384)	(10.158)
H	5.54	64.17	57.004	78.045
	(18.81)	(36.30)	(7.3568)	(7.03)
I	13.83	29.65	78.048	18.824
	(11.10)	(6.61)	(20.811)	(9.1563)
J	0.83	-0.53	-26.785	17.531
	(1.20)	(0.78)	(35.647)	(16.403)
K	33.43	18.85	35.74	27.166
	(23.52)	(10.01)	(20.393)	(6.6007)
L	25.07	10.00	6.8141	21.698
	(4.71)	(6.41)	(35.144)	(18.79)
M	27.80	121.28	19.064	-17.171
	(40.50)	(57.60)	(26.851)	(157.22)
N	48.41	43.90	23.327	55.884
	(46.51)	(20.38)	(45.913)	(32.101)

---

[a] Corresponding to recover parameter values summarized in Table 2.

[b] Corresponding to GMM estimates from Table 5.

Table 8: Observed vs. first-best total wage bill implied by estimates

	First approach			Second approach[a]		
RETAIL CHAIN	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	1.49	1.13	0.24	1.808	1.301 (0.018)	0.28
B	4.35	3.16	0.27	4.356	3.160 (0.104)	0.27
C	1.33	1.02	0.23	3.347	2.498 (0.337)	0.25
D	0.39	0.31	0.21	1.349	0.836 (0.239)	0.38
E	20.62	15.96	0.23	30.344	23.793 (0.934)	0.22
F	—	—	—	5.286	3.872 (0.055)	0.27
G	0.91	0.68	0.25	5.893	4.239 (0.109)	0.28
H	2.91	2.27	0.22	19.318	17.222 (0.197)	0.11
I	58.95	44.84	0.24	178.431	137.952 (1.785)	0.23
J	1.60	0.84	0.48	9.296	5.344 (0.254)	0.43
K	9.29	7.46	0.20	14.977	11.815 (0.192)	0.21
L	0.73	0.50	0.32	6.939	5.044 (0.224)	0.27
M	32.75	26.32	0.20	58.289	41.084 (3.123)	0.30
N	3.41	2.77	0.19	3.343	2.714 (0.324)	0.19

---

[a] Corresponding to GMM estimates from Table 5.

[b] Note 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 9: 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	3588.8	207.81	1.107	1.853
		(506.02)	(2106.6)	(504.02)	(0.238)	(0.603)
B	32	1822.8	3733.1	1861.4	1.822	1.887
		(649.07)	(5201.8)	(6434.5)	(0.276)	(0.596)
G	24[c]	3614.6	2448.2	691.47	0.918	2.018
		(772.51)	(10047)	(1387.5)	(0.056)	(0.569)
M	60	3057.6	5681.9	2549.3	1.128	1.730
		(1497.4)	(4339.5)	(23681)	(0.187)	(0.630)

---

[a] Exponent on $h(\cdot)$, link function specification.

[b] asymptotically distributed $\chi^2(2)$ under null that the moment conditions in Eq. (21) hold.

[c] Fewer 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 10: Average percentage of wage differentials accounted for by effort differentials: incorporating turnover rates

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

---

[a] Corresponding to GMM estimates from Table 9.

Table 11: Observed vs. first-best total wage bill implied by estimates: incorporating turnover rates

	Second approach[a]		
RETAIL CHAIN	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

---

[a] Corresponding to GMM estimates from Table 9.

[b] Note 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.