MANAGEMENT SCIENCE
Vol. 55, No. 8, August 2009, pp. 1306–1320
ISSN 0025-1909 | EISSN 1526-5501 | 09 | 5508 | 1306

informs

DOI 10.1287/mnsc.1090.1021

Relative Performance Compensation, Contests, and Dynamic Incentives

Pablo Casas-Arce
Department of Economics and Business, Universitat Pompeu Fabra, and Barcelona Graduate School of Economics and IESE, 08005 Barcelona, Spain, pablo.casas@upf.edu

F. Asís Martínez-Jerez
Harvard Business School, Harvard University, Boston, Massachusetts 02163, asis@hbs.edu

Contests (or tournaments) are pervasive in organizations. They help performance evaluation by eliminating common shocks affecting agents' performance. However, tournaments are less effective when participants have heterogeneous ability because participants may conclude that the ability gap is too large to be overcome by their effort. Our theoretical analysis shows that a similar loss of motivation arises when tournaments take place over multiple periods because interim performance acts in a way that is similar to heterogeneous ability. Analyzing the sales contests organized by a commodities company, we document that winning participants decrease their effort as their lead extends, whereas the effort of trailing participants fades only when the gap to a winning position is very large. We also show that, on average, when contests are introduced they induce a higher level of effort among participants, although the incentives weaken as the number of participants increases. Finally, we demonstrate that although retailers respond to the multiple performance dimensions of the incentive program in part by shifting effort toward sales of more expensive products, they channel most of the increased effort toward reaching more customers.

Key words: tournaments; relative performance compensation; dynamic incentives; multitasking
History: Received October 23, 2006; accepted December 20, 2008, by Stefan Reichelstein, accounting. Published online in *Articles in Advance* June 1, 2009.

1. Introduction

Relative performance compensation schemes are pervasive in organizations. From sports competitions to sales contests and job promotions, we observe individuals and teams being measured against one another in pursuit of a reward. In this paper, we study the impact of several tournament features on contestants' performance using data from a commodities manufacturer that organized several sales contests among its retailers. We also provide a theoretical analysis that sheds light on the behavior of participants in dynamic tournaments such as the one in our study.

There is an extensive body of theoretical work on tournaments and relative performance evaluation (RPE), beginning with the pioneering work of Lazear and Rosen (1981) and Holmstrom (1982). The main theme of this literature relates to the role of RPE incentive schemes in eliminating common shocks that affect agents' performance, hence reducing the cost of providing incentives. Although these studies have been limited to static settings in which homogeneous participants reach a symmetric equilibrium,[1] many contests are designed as multiperiod schemes in which interim performance generates heterogeneity among contestants.[2] When prizes are allocated on the basis of total performance over several periods, participants have to outperform any advantage gained by a competitor in the past in order to win. Sports competitions offer many such examples; for instance,

[1] An exception is Meyer (1992), who considers a biased dynamic tournament. In her model, participants are heterogeneous. Nonetheless, her assumptions (namely the presence of only two participants, and the symmetry in the distribution of the shocks) guarantee that the equilibrium is still symmetric.

[2] Recently, some progress has been made in the study of asymmetric tournaments by modeling them as all-pay auctions (see, for instance, Moldovanu and Sela 2001). These models incorporate differences in the marginal cost of effort. The emphasis in this paper, instead, is on the heterogeneity that arises in dynamic tournaments. Our setup also differs from the papers that study elimination tournaments (most notably, Rosen 1986). In these models, effort in a given period only affects the probability of winning in the current period, whereas in a dynamic tournament, effort in any given period affects the probability of winning at the end of the contest. Hence effort does not depend on past performance in an elimination tournament, but it does in a dynamic one. There is also related work on dynamic games, such as the war of attrition or research and development races. However, the payoffs of these games are different from those of a tournament. The end of a war of attrition or a race occurs endogenously when all contestants but one drop out of the war or when an innovation takes place. Yet in a tournament, the end arrives exogeneously, after a prespecified lapse of time.

in cycling, the Tour de France takes place over three weeks. But economic contests also have dynamic aspects. Promotion decisions often take place over several years, and battles for standards adoption (e.g., operating systems, VCRs, HDTV) span several cycles.

Our research setting, a contest among the retailers of a commodities manufacturer,[3] is another instance of multiperiod contest dynamics. It provides an opportunity to analyze the consequences of different design choices on participant incentives. In this paper, we make progress toward understanding the heterogeneity that arises from the dynamic nature of a multiperiod contest, the size of the tournament, the imposition of qualification requirements, or the weighting of the different dimensions of performance (such as upselling or reaching new customers).[4]

We are interested in four questions regarding the implementation of contest-based incentive schemes: Are participants motivated to exert more effort after the introduction of such a scheme? If so, how does the size of the contest affect incentives? When the contest takes place over several periods, how do incentives evolve as the tournament progresses? And finally, how are these considerations affected by the presence of multitasking—are the results observed due to an increased effort to expand the client portfolio or to upsell clients?

We find evidence suggesting that the introduction of the contest significantly increased retailers' effort. This is consistent with earlier work documenting significant improvements in output after the introduction of piece rates (Lazear 2000). Moreover, we also show that the effort exerted by each participant is inversely related to the size of the contest (even while keeping constant the number of prizes per participant), which is consistent with the theoretical results obtained by Gibbs (1996).

We exploit the features of our site to analyze the dynamic implications of tournament incentive schemes that have been overlooked by the literature. We find that having a strong lead in the tournament attenuates the incentive effect, as our theory suggests. On the other hand, retailers that are trailing in the rankings increase their effort in trying to catch up, losing motivation only when the performance gap with winners is very large. These results document important dynamic effects—a reflection of the nonlinearities of these incentive schemes (see Holmstrom and Milgrom 1987)—that should be taken into account.

Finally, we study the influence of the performance measure used in the contest on the allocation of effort across tasks. Retailers were responsible for distributing three different types of products, each carrying a different weight in the computation of performance.[5] By comparing the changes in volume and mix of products sold, we show that the performance increases are mainly due to an increase in the effort to reach more clients, although we also observe an increase in the effort to upsell all clients to higher value-added products.

Our paper contributes to the literature on RPE by gauging (both theoretically and empirically) the incentive effects of the heterogeneity induced by the dynamic nature of multiperiod contests. Our results suggest that contestant effort depends on the interim ranking, with effort being highest for those in intermediate positions and lowest for those at the extremes of performance. We also contribute to the literature on tournament design by examining the incentive effects of the number of participants in the contest. In contrast with other studies (Orrison et al. 2004), we find that increasing the contest size reduces contestants' effort. Finally, we add to the mounting body of evidence that suggests that performance improves following the introduction of an RPE incentive scheme. We also extend the existing literature by demonstrating the extent to which changes in performance under an incentive plan can be attributed to changes in effort allocation in a multitasking setting.[6] The last point raises the question of the relative cost of changing customers' behavior versus reaching more customers for salespeople in this setting.

The structure of this paper is as follows: §2 provides the institutional background, §3 discusses a simple model of multistage tournaments, §4 discusses the empirical approach, §5 describes the results of the empirical analyses, and §6 concludes.

[3] Our confidentiality agreement prohibits disclosure of additional details about the company.

[4] The initial empirical evidence on relative performance compensation has largely focused on showing the consistency of observed compensation/promotion schemes with theoretical predictions (e.g., Antle and Smith 1986, Gibbons and Murphy 1990). The first attempts to analyze contest design elements—notably the number and size of the prizes—mainly used experimental settings (e.g., Bull et al. 1987, Orrison et al. 2004, Hannan et al. 2008) or evidence drawn from sports competitions (e.g., Ehrenberg and Bognanno 1990). Until recently, there was scant evidence on the incentive effects of RPE systems using actual business settings (e.g., Knoeber and Thurman 1994, Bandiera et al. 2005, Matsumura and Shin 2006). However, none of these papers use a tournament setting (but rather various forms of RPE), and none consider dynamic effects. There is also some evidence on promotion tournaments (e.g., Main et al. 1993), but dynamic considerations are absent from these papers as well.

[5] The company put more weight on certain products in the performance index used to rank participants in order to induce the retailers to increase their efforts to sell them.

[6] For evidence on multitasking in nontournament settings, see Slade (1996) and Brickley and Zimmerman (2001).

2. Institutional Setting

2.1. Market

The firm we study operates in a commodity market in a less-developed country. The national economy has grown systematically in the last decade albeit with high volatility. The country's inflation, unemployment, and political uncertainty are on par with most other underdeveloped economies, making forecasting difficult.

Manufacturers in this industry either sell directly to big consumers and wholesalers or use a network of independent retailers that also provides product advisory and financial services to small and financially weak consumers. Retailers typically carry a full line of products that complement the manufacturers' offerings and may carry more than one brand of each product.

As in most commodity markets, input cost fluctuations determine price evolution. Competitors closely track and mirror each other's prices. The firm, in a process that is typical of the industry, sets prices weekly unless major changes in market conditions dictate a shorter revision period. Retailers, in turn, set prices for the end consumer by adding a small margin over input costs. Retailers generally order merchandise from the firm twice a week.

In this market, sales are final and product returns are not allowed. The products are bulky, are expensive to store, and have a relatively short shelf life due to both physical degradation and customers' desire for fresh merchandise, an attitude that the firm promotes. In fact, the average time elapsed from the moment the product leaves the production line until it is sold to the end consumer is eight days. Expensive storage, limited shelf life, and volatile prices limit the ability of salespeople and retailers to window-dress their performance by selectively choosing the timing of their sales.

The end consumer's prior decision to undertake a project determines the decision to buy the product and the quantity to buy. Thus retailers may not impact the quantity consumed by the consumer but may influence consumer choice of quality level. In this sense demand is similar to that for automobile tires: A retailer may influence the quality of the tires but will not be able to sell more than four tires to the same customer. Thus, as we confirmed in our interviews with retailers, to increase sales volume retailers needed to reach a larger number of potential customers.

2.2. Tournaments

The firm was acquired by a foreign multinational in 2000. After restructuring the manufacturing division, the new managers shifted their focus to the commercial strategy. The new commercial strategy focused on decommoditization of the firm's products. Product differentiation and alliances with small and medium retailers were the main initiatives to achieve this goal.

The product strategy involved manufacturing improvements and a customer education campaign. The restructured manufacturing facilities quickly implemented more stringent quality standards and facilitated the introduction of new product offerings. As a result, the firm is now considered best-in-class and commands a small premium over its peers. The firm's main product is almost identical to that of its competitors. It also sells two specialty products (premium and enhanced) that bundle the base product with other components. This differentiated offering commands a price premium over the main product.

The communication strategy stressed freshness and physical attributes of the product that could be easily tested by the end consumer. Retailers were also instrumental in the education campaign and were required to complete a basic training program that prepared them to advise consumers on the optimal use of the firm's products.

The company decided to focus on small and medium retailers with the objective of avoiding the price pressure of big traders. This approach limited the bargaining power of any single retailer and the interest of other manufacturers in its business. To secure retailers' loyalty, the firm instituted a volume-discount program, similar to a bonus bank,[7] in which the firm paid retailers a rebate in semiannual installments contingent on the retailer continuing to sell the firm's products.[8]

After observing performance for more than a year, the firm managers believed that the retailers' margins and loyalty discounts did not provide sufficient incentives to achieve the target growth rates. To stimulate sales they decided to try a sales tournament, as market uncertainty created an ideal setting for the use of an incentive mechanism that eliminated the performance impact of common shocks (Lazear and Rosen 1981).[9] Moreover, in the view of the firm's management, the use of prizes instead of additional volume

[7] A bonus bank is a compensation scheme in which the payment of a bonus is deferred over several periods to achieve consistent multiperiod performance.

[8] The volume discount program was initiated in January 2001 and remained unchanged throughout our sample period. The tournament was the only change in the incentive system during the period of analysis. The bonus is calculated monthly and consists of a discount of four cents for each of the first 400 units, three cents for each of the units between 400–800, two cents for each of the units between 800 and 1,200, and one cent for each of the units above 1,200.

[9] During interviews firm managers asserted that the introduction of the new incentive system was not motivated by changes in the operating or competitive environment but simply by the perceived need to increase retailers' efforts to sell the firm's products.

discounts would reduce the probability of retailers passing their incentives to end consumers and triggering a price war.[10]

In May 2002, the firm designed a set of tournaments among retailers called the Reward Program. Its goal was to motivate retailers to increase overall sales as well as the proportion of premium products sold and to reduce variability in sales volume (which would simplify scheduling for the manufacturing department).[11] Participation in 2002 was optional, and almost 70% of the firm's retailers signed up for the contest, giving a total of 500 participants. All the retailers were enrolled in the contest for the second season of the Reward Program in 2003.

Except for the first edition, which ran from May to December 2002, the tournaments ran throughout the calendar year. The program consisted of eight different contests, with each contest corresponding to one of the eight regions in which the company divided the market. Different factors impacted economic performance in each region, and the firm reasoned that independent contests would help eliminate the impact of region-specific shocks on retailer performance.

The total number of winners was 50 (51 in 2003). All winners, regardless of their final ranking, received the monetary equivalent of $3,000, about 50% of the annual profit of the median retailer. At the start of the tournament, each region was allocated a fixed number of prizes, with approximately one prize for every ten participating retailers in 2002 (see Table 1).

The firm set individual monthly sales targets for each retailer.[12] These targets leveled the playing field by setting higher goals for larger retailers, similar to the handicap system of amateur golf championships. To avoid being eliminated from the contest, each month a retailer was required to buy from the firm at least 150 units of product and 70% of the target quantity for the month.[13] Performance was evaluated using a scoring system that awarded points for each unit sold. The points per unit varied by product line: sales of basic, premium, and enhanced products earned one, five, and nine points per unit, respectively. Rankings were calculated within each region based on the cumulative percentage of weighted sales (in points) over target (also in points) for the year.

At the start of the contest, each retailer was informed of the program rules, the number of prizes and participants in its region, and its individual monthly targets (but not its competitors'). On a monthly basis, performance was assessed and each retailer was informed of its own interim ranking but not of the number of remaining competitors or their scores.

2.3. Descriptive Statistics

As Table 1 shows, the eight regional tournaments were heterogeneous in size, both by number of contestants (ranging from less than 10 to more than 100) and by sales volume (ranging from 5,000 to 70,000 units per month). Firm sales reflect an overall market contraction in 2003, though steep declines in some regions contrast with market growth in others. The magnitude and regional diversity of the market shocks suggest that the choice of multiple tournaments rather than a single, overarching contest was appropriate in this environment. The firm's basic product accounted for the overwhelming majority of its sales by weight (93%). Its premium and enhanced products represented 1% and 6% of sales, respectively.

Panel B of Table 1 shows the evolution of drops (contestants not satisfying the qualification requirements) during the tournament. All regions follow a similar pattern, with most of the attrition occurring within the first three months of the campaign. The absence of drops in May 2002 reflects the firm's policy of waiving the qualification requirements during the program's first month.

Figures 1 and 2 show the distribution of monthly performance (sales over target) and raw sales of the basic product for all tournament participants. Consistent with the above discussion, both distributions are fairly smooth and do not show anomalous behavior around critical points (the 70% qualification requirement for sales over target or the thresholds for the volume discount), which might have been indicative of window-dressing activities.

In summary, the heterogeneous starting characteristics of the eight tournaments and their evolution over time generated a diverse set of observations with which to test tournament theories.

3. Simple Model of a Dynamic Tournament

This paper aims to gauge the effect of the dynamic nature of multiperiod tournaments on incentives. As

[10] The firm managers thought that if they lowered the price to retailers—for instance by using piece rates—retailers would respond by passing along some of this price reduction to the end consumer to benefit from higher sales volume at similar margins, potentially triggering a price war. However, if the reward was made contingent on exceeding the performance of other retailers, the uncertainty of attaining the prize might deter a retailer from sharing the expected cost reduction with the end consumer. A formal analysis of this argument is beyond the objectives of this paper.

[11] Our interviews with the managers of the firm indicate that the last objective was secondary and ultimately unsuccessful. For these reasons, we do not analyze it here.

[12] These targets were set for the sole purpose of evaluating performance in the tournament and did not affect any other element of the retailers' incentives.

[13] During the first month of the 2002 tournament, the firm decided not to drop any participant that failed to comply with the participation requirements. Thereafter the tournament proceeded according to these rules.

Table 1 Summary Statistics

Panel A: Descriptive statistics

Region	Initial number of contestants		Contestants in final round		Number of winners	Average monthly sales (units)			Breakdown of sales (by weight)[b] (%)		
						Pretournament	Posttournament				
	2002	2003	2002	2003		2002	2002	2003	Basic	Premium	Enhanced
A	100	165	13	47	10	68,270	61,913	63,425	95	3	2
B	21	56	5	6	2	13,655	15,125	13,191	96	0	4
C	65	73	6	7	4[a]	40,710	44,976	33,969	78	0	22
D	64	78	6	12	6	27,805	28,097	26,396	97	0	3
E	6	7	5	4	1	4,909	4,878	5,526	92	0	8
F	105	148	38	23	14	62,174	65,540	44,917	97	1	2
G	87	119	10	43	8	47,774	51,802	64,681	95	2	3
H	52	105	8	9	5	72,561	81,990	53,312	93	0	7
Total	500	751	91	151	50	337,858	354,321	305,417	93	1	6

Panel B: Retailers dropping from the tournament by month

Month	2002		2003	
	Drops (#)	Drops (%)	Drops (#)	Drops (%)
January			282	47
February			96	16
March			40	7
April			81	13
May	0	0	45	7
June	136	32	5	1
July	205	49	13	2
August	29	7	6	1
September	13	3	3	0
October	11	3	26	4
November	15	4	3	0
December	10	2	3	0
Total	419	100	603	100

Notes. Panel A describes the characteristics of the eight tournaments organized by the firm among its retailers. The tournaments ran from May to December 2002 and from January to December 2003. Each winner received \$3,000 cash, regardless of the region or final rank achieved in the tournament. Monthly targets were set for each retailer in advance. Rankings are calculated within each region based on the percentage of realization over target, provided that certain requisites are met. Panel B displays the number of retailers dropping out of the tournament by month. Monthly targets were set for each retailer in advance. To stay in the tournament a retailer had to buy at least 70% of its target and a minimum of 150 units of product. Retailers not satisfying both requirements were dropped from the tournament. Retailers in panel B may have violated one or both conditions.

[a]In 2003 the number of winners increased to five for region C.

[b]Sales to all retailers participating in the tournament. Percentages are similar for the region as a whole.

a tournament evolves, the interim rankings may affect contestant effort. Those that fall behind need to exert more effort to climb the rankings and qualify for a prize. With a sufficiently low ranking, they may reduce effort or stop trying (Müller and Schotter 2009).[14] Conversely, contestants at the top might see little threat of losing and therefore decrease their effort. Consistent with the expectancy-valence theory (Vroom 1964), motivation should be highest for those whose change in position could affect the outcome of the tournament. In this section, we investigate the robustness of this intuition with a stylized two-period model that incorporates features of our empirical application. We conjecture that the conclusions extend to a multiperiod contest such as our tournament setting that spans multiple months.

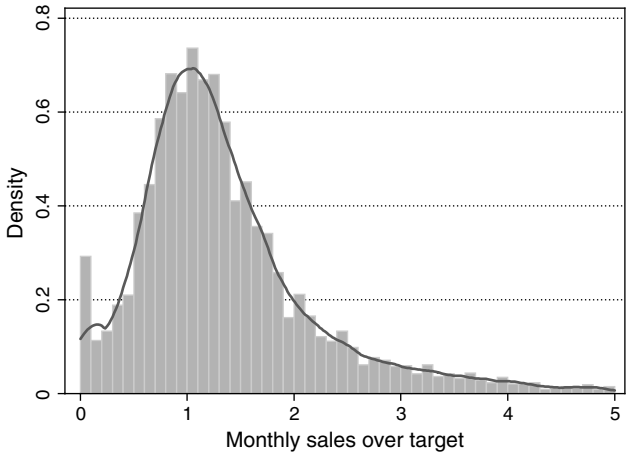
3.1. Basic Framework

Consider a tournament with an infinite number of identical participants.[15] A fraction μ of them will win

[14] The intuition behind this argument is very similar to that for stock options. When options are under water, the manager finds herself in the flat part of the incentive scheme and hence loses motivation (Murphy 1999).

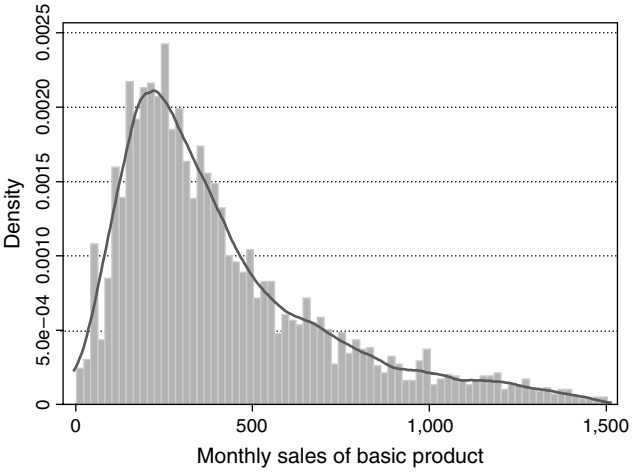
[15] The assumption of an infinite number of participants simplifies the analysis by avoiding the complicated combinatorial calculations that arise in finite tournaments. In a working paper (Casas-Arce and Martínez-Jerez 2009), we have solved a tournament with heterogeneous and finite participants using the framework of an all-pay auction with handicaps. Although in this approach we make

Figure 1 Distribution of Monthly Sales Over Target of Tournament Participants



Note. The line shows the estimated density.

Figure 2 Distribution of Monthly Sales of Basic Product of Tournament Participants



Note. The line shows the estimated density.

a prize, where all the prizes are identical and have a value that we normalize to 1. The prizes are awarded to the best performers on the basis of the output produced over two periods, $t = 1, 2$. Each participant i produces $\tilde{y}_{it} = e_{it} + \tilde{u}_{it}$, where $\tilde{y}_{it}$ is a measure of performance or sales, e_{it} is effort in increasing sales, and $\tilde{u}_{it}$ is a random shock representing anything beyond the control of the retailer.[16] Such uncertainty can be common to all retailers (e.g., regionwide economic conditions) or individual and specific (such as when a potential customer receives a cash windfall and makes an unexpected purchase). The first type should affect all retailers equally and, because of its relative nature, tournament performance should be insensitive to it. For this reason, we neglect this type of uncertainty and assume that the $\tilde{u}_{it}$ are independent and identically distributed (i.i.d.), with a differentiable pdf ψ and cdf Ψ. Effort is incurred at a cost of $c(e_{it})$, which satisfies $c'(0) = 0$ and has increasing marginal costs, i.e., $c''(e) > 0$.[17]

At $t = 1$, all participants are identical, and hence they reach a symmetric equilibrium anticipating what will happen in the future. We focus here on the outcome of the last period. At the end of period 2 there will be a distribution of realized performances $y_{i1} + y_{i2}$ that will determine a threshold $\bar{y}_\mu$ such that player i wins if and only if $y_{i1} + y_{i2} \geq \bar{y}_\mu$. The threshold is endogenously determined so that the fraction of participants who obtain a prize is exactly equal to μ.[18] The players can predict this threshold because, by the law of large numbers, the empirical distribution of shocks coincides with the theoretical distribution.[19] Foreseeing the performance it will need to achieve to win, $\bar{y}_\mu$, the objective at time $t = 2$ of a participant that obtained output y_{i1} in period 1 is to maximize the difference between the probability of winning a prize and the cost of the effort: $\Pr(y_{i1} + \tilde{y}_{i2} \geq \bar{y}_\mu) - c(e_{i2})$, where $\Pr(y_{i1} + \tilde{y}_{i2} \geq \bar{y}_\mu) = 1 - \Psi(\bar{y}_\mu - y_{i1} - e_{i2})$. The solution to this problem will have to equate the marginal cost of exerting effort to the marginal benefit of an increased probability of winning, due to the higher sales:

$$\psi(\bar{y}_\mu - y_{i1} - e^*_{i2}) = c'(e^*_{i2}).$$

The marginal benefit is given by the effect of effort on the probability of beating the threshold and winning a prize. The equation then defines the optimal effort level e^*_{i2} as a function of past performance y_{i1}.[20] If we further assume that the distribution of the shocks is unimodal—an assumption that implies that extreme levels of luck (both good and bad) are rare—we can state that:[21]

PROPOSITION 1. *Suppose that ψ is unimodal; then effort $e^*_{i2}(y_{i1})$ has a $\cap$-shape.*

some restrictive simplifying assumptions—namely, the absence of uncertainty and linearity in the cost of effort—our paper is one piece of evidence that allows us to think that our result is robust to having a finite number of participants.

[16] We use $\sim$ to distinguish random variables from outcomes.

[17] If we think of effort as the number of working hours, it is harder to work the eleventh than the seventh hour.

[18] Notice that the threshold is an outcome of the equilibrium because the realized output depends on the distribution of effort in the population. Moreover, given the equilibrium effort choices, the threshold is deterministic.

[19] This follows from our assumption of an infinite number of participants (see Al-Najjar 2008 for a formal treatment).

[20] Note that we can allow for the coexistence of a piece-rate compensation scheme ($w_{it} = \beta y_{it}$). The equilibrium effort would now be higher and satisfy $\beta + \psi(\bar{y}_\mu - y_{i1} - e^*_{i2}) = c'(e^*_{i2})$. However, the same results would still hold.

[21] This assumption, while restrictive, is satisfied by most common distributions, such as the Gaussian.

Proof. We provide here a proof assuming a concave objective function.[22] In this case, the maximization problem has a unique solution characterized by the above first-order condition, and by the implicit function, theorem $e^*_{i2}(y_{i1})$ is differentiable and satisfies $\partial e^*_{i2}/\partial y_{i1} = -\psi'(\bar{y}_\mu - y_{i1} - e^*_{i2})/(\psi'(\bar{y}_\mu - y_{i1} - e^*_{i2}) + c''(e^*_{i2})) > -1$. For a concave objective function the denominator is positive, and therefore, the sign of $\partial e^*_{i2}/\partial y_{i1}$ is the opposite of the sign of ψ'. Furthermore, $y_{i1} + e^*_{i2}$ is increasing in y_{i1}. Hence there exists a constant $\hat{y}_{i1}$ such that effort is increasing in past output if $y_{i1} \le \hat{y}_{i1}$ and decreasing otherwise. □

Participants with a strong initial lead (high y_{i1}) realize that only a disastrous shock could cause them to lose. Because the probability of such a shock is very small, and effort is not very useful in trying to avoid such a scenario, these participants will tend to put out lower effort in the following months. Similarly, those that are clearly trailing (low y_{i1}) realize they can only win if they are very lucky in the last period, which is quite unlikely. Effort cannot help them much either, and they will also tend to decrease their efforts. In contrast, participants in intermediate positions are more motivated because their effort can affect whether they win or lose.

Our empirical setting contains a multitasking element because the performance measure includes sales of three different commodities. Adding this multitasking element should not alter any of the results above. Moreover, we should expect all the dimensions of the multitasking problem to respond in the same direction, with greater emphasis on the dimension that is more sensitive to the retailer's effort (Holmstrom and Milgrom 1991, Feltham and Xie 1994).

3.2. Introduction of Participation Constraint

We want to assess how the participation constraint of the Reward Program alters the previous results. We can model this scenario by assuming that there is a threshold $\underline{y}$ such that a participant is dropped if $y_{it} < \underline{y}$ in any given period (including the last one).[23] Under the new rules, a participant that reaches the second period wins a prize as long as both $y_{i2} \ge \underline{y}$ and $y_{i1} + y_{i2} \ge \bar{y}_\mu$, where $\bar{y}_\mu$ is a new threshold endogenously chosen for the new distribution of realized performance among the survivors. Notice that there is only one binding constraint: $y_{i1} + y_{i2} \ge \bar{y}_\mu$ whenever $y_{i1} < \bar{y}_\mu - \underline{y}$, and $y_{i2} \ge \underline{y}$ otherwise. For the participants that are lagging, avoiding disqualification is not enough to win a prize, and therefore they focus on catching up. In contrast, the top contestants are virtually guaranteed a prize if they avoid a drop, and hence their motivation is driven by the fear of being dropped rather than of falling behind (i.e., $y_{i2} \ge \underline{y}$ is binding). A different way to see this is by noticing that dropping out must be more costly for highly ranked participants (they have a higher probability of winning, and consequently their option value of remaining in the tournament is larger). Hence, the participation constraint must induce more effort among them. Therefore, this feature will make the ∩-shape in Proposition 1 less pronounced.

3.3. Participants with Heterogeneous Ability

So far we have assumed that participants are homogeneous. Now we relax this assumption and allow for participants to have heterogeneous abilities. This may happen in our setting, for instance, if the handicap system of the contest is imperfectly designed and does not level the playing field. In this case, it is reasonable to assume that retailers (because they do not know their competitors' targets) are unaware of their relative advantage. However, they may learn about this advantage over time as they acquire information about their rankings. We can model this by assuming that output is now $\tilde{y}_{it} = e_{it} + \tilde{a}_i + \tilde{u}_{it}$, where $\tilde{a}_i$ is player i's ability or residual advantage. There are now two sources of uncertainty for a contestant: its luck in selling the products, $\tilde{u}_{it}$, and its targets relative to those of its competitors, $\tilde{a}_i$. To model the learning process, we can make some simplifying assumptions, namely that $\tilde{a}_i$ and $\tilde{u}_{it}$ are normally distributed, $\tilde{u}_{it} \sim N(0, \sigma_u^2)$, $\tilde{a}_i \sim N(0, \sigma_a^2)$ (the assumption of zero means is simply a normalization). All participants are assumed to know the underlying parameters of the model. After having exerted effort e_{i1} and obtained output y_{i1}, player i will update its understanding of its relative advantage. Its posterior belief will have an expectation of $\alpha(y_{i1} - e_{i1})$, where $\alpha = \sigma_a^2/(\sigma_u^2 + \sigma_a^2)$ represents the information content of output about a_i. The higher the first period output, the higher the expectation that the tournament favors contestant i. And furthermore, this belief is reinforced when the initial performance is very informative about this advantage (high α). The new first-order condition becomes $\phi(\bar{y}_\mu - e^*_{i2} - y_{i1} - \alpha(y_{i1} - e_{i1})) = c'(e^*_{i2})$, where ϕ denotes the pdf of a $N(0, \Sigma)$ and

[22] A general proof is available from the authors upon request. A sufficient condition for having a concave objective function is that $c''(e) > |\psi'(y)|$ for all e and y, which means that the cost of the marginal effort grows at a faster pace than its impact on the probability of winning. Although concavity over the whole domain of the function may seem restrictive, we keep the assumption for tractability and expositional purposes because it greatly simplifies the proof of Proposition 1. If the objective function is not concave, the solution $e^*_{i2}(y_{i1})$ may not be continuous. Nevertheless, the same results would still hold.

[23] In this discussion we assume that $\underline{y}$ and $\bar{y}_\mu$ are measured in the same units. The rationale here would not be completely applicable to the 150 unit constraint because the program performance metric is percent sales over target rather than unit sales. However, we think that the analysis of this section is appropriate because the 150 unit constraint is the cause of only 7% of the drops.

$\Sigma = \sigma_u^2 + \sigma_a^2 \cdot \sigma_u^2/(\sigma_a^2 + \sigma_u^2)$. This first-order condition is analogous to the one we had above, and hence the result in Proposition 1 is robust to introducing heterogeneity in ability. The main difference comes from the term $\alpha(y_{i1} - e_{i1})$. When a contestant obtains an output y_{i1}, it will expect this performance to persist in the second period, and $\alpha(y_{i1} - e_{i1})$ captures this persistence. Those who did well in the past (high y_{i1}) believe they are favored by the target-setting process and reduce effort as if they had done much better (by a factor of α). Likewise, those who did badly (low y_{i1}) respond in the second period as if they had done much worse. Hence the ∩-shape is exaggerated with heterogeneity because confidence of leaders and demotivation of trailers arise earlier. Moreover, if we introduce attrition in the model along with ability, the threshold will be less likely to bind participants who benefited from the target setting process, diminishing the threat of dropping, and thus exacerbating the ∩-shape.

4. Empirical Approach

For the empirical implementation, we consider the following theoretical model for retailer output:

$$output_{i,m} = effort_{i,m} + Z'_{i,m}\gamma + a_i + u_{i,m},$$
$$\text{where } effort_{i,m} = X'_{i,m}\beta. \quad (1)$$

Output is a function of the effort retailer i makes during month m, where this effort is affected by the variables in $X_{i,m}$, and β measures the size of their effect. We also allow for some control variables $Z'_{i,m}$, and the existence of a retailer fixed effect, a_i, which could be interpreted as either the size, ability, or any other intrinsic (time invariant) characteristic of the retailer that affects its performance in the tournament. And finally, there is an idiosyncratic component, $u_{i,m}$.

We are interested in the determinants of effort. First of all, $effort_{i,m}$ should depend on whether the retailer is participating in the tournament during month m (this is the incentive effect of the introduction of the program). Furthermore, once in the tournament, the retailer's effort level should depend on its relative position at the start of the period (which affects the probability of winning the prize, and hence incentives).

To test whether the tournament has an effect on incentives, we use the full sample to compare the sales of participants to nonparticipants using a difference-in-differences estimation. The *Diff-in-Diffs* coefficient compares the change in output levels before and after the introduction of the tournament for participants versus nonparticipants. A positive coefficient would indicate that participants increase output during the months of the tournament by more than nonparticipants do. This will be taken as a measure of the effect of the Reward Program. We will also include additional controls ($Z_{i,m}$) in some specifications, such as time fixed effects.

When analyzing the dynamic incentives of the tournament, we limit the sample to those observations in which the retailer is currently participating in the competition. To understand the effects of the retailer's ranking on its willingness to exert effort, we generate the variables $distance\ for\ winners_{i,m} = \max\{0, prizes - ranking_{i,m}\}$ and $distance\ for\ trailers_{i,m} = \max\{0, ranking_{i,m} - prizes\}$.[24] The first of these, $distance\ for\ winners_{i,m}$, is the maximum number of places a winning retailer could fall and still remain in the winners' circle; when the retailer is not among the winners, this measure is zero. $Distance\ for\ trailers_{i,m}$ is the number of positions a nonwinning retailer would have to climb to become a winner; this measure is zero when the retailer is winning. Finally, we also include the *Lagged Output* as an explanatory variable in some specifications, to control for the possibility of autocorrelation in the shocks to performance.

The inclusion of predetermined or lagged dependent variables in a regression yields biased estimates when using fixed effects to control for individual heterogeneity (see Nickell 1981).[25] To obtain consistent estimates we use a method proposed by Anderson and Hsiao (1982) and refined by Arellano and Bond (1991). It involves differencing Equation (1) to remove the fixed effects:

$$\Delta output_{i,m} = \Delta X'_{i,m}\beta + \Delta Z'_{i,m}\gamma + \Delta u_{i,m}.$$

Using ordinary least squares (OLS) on the differenced equation would yield biased estimates if $X_{i,m}$ or $Z_{i,m}$ contain predetermined or lagged dependent variables. However, we can use lagged values of these variables as instruments. The two approaches differ only in the use of these lags. Whereas Anderson and Hsiao use only the first lag, Arellano and Bond use an efficient GMM estimator containing all available lags. This method assumes that errors, $u_{i,m}$, are i.i.d. Following Arellano and Bond (1991), we check the validity of this assumption by testing for the existence of first-order negative autocorrelation, and the absence of second-order autocorrelation in the terms $\Delta u_{i,m}$.[26]

[24] Although the theory presented here does not precisely identify the retailer with the highest effort level, we decided to distinguish the effect for winners and trailers for practical purposes.

[25] Notice that the distance measures are predetermined. They will not be correlated with the contemporaneous shocks if $u_{i,m}$ are i.i.d. However, the distance measures are generated from previous outputs and hence correlated with past shocks.

[26] Differenced errors are always autocorrelated. But if the errors are autocorrelated, we would also observe second-order autocorrelation in the differenced errors.

5. Empirical Results

5.1. Effects of Introduction of the Tournament

In this section, we test the incentive effects of the introduction of the Reward Program. Specifically, we analyze three dimensions: (i) whether participation in the tournament leads retailers to increase their effort; (ii) whether the incentive effect varies with the size of the tournament (as measured by the number of participants); and (iii) whether retailers channel their effort increase through an increase in the number of clients reached or through an upselling of customers to more value-added products (in other words, what types of trade-offs retailers make in response to the multitasking component of the Reward Program). As in any field study, we cannot rule out with certainty the possibility that the observed results are the consequence of an exogenous shock that triggered the adoption of the Reward Program. However, the use of a control group in our estimations and our field interviews seem to indicate that we are indeed observing the impact of a change in the incentive system. Furthermore, none of the firm competitors adopted a similar compensation system before or at the same time, suggesting that the Reward Program was not a response to an external shock.

We use evidence from difference-in-differences estimation to test whether retailers increase their effort after the introduction of the Reward Program. Ideally, we would prefer to test this hypothesis using a sample of retailers that were randomly assigned to the program. Unfortunately, the firm offered every retailer the option to join. Lack of adequate retailer information to run a selection model limits our ability to correct for potential biases. However, interviews with retailers revealed that the most common reason for refusing contest participation was distaste for signing the document explaining the program rules, suggesting that self-selection was not influenced by the retailers' expected performance. This observation gives more power to the use of nonparticipants as a control population. Moreover, we attempt to empirically address our concern with self-selection in two ways. First, we include fixed effects in our specifications. This should control for any unobserved characteristics that are correlated with the choice to join or abstain from the tournament. And second, we analyze whether the incentive effects of participating in the 2003 tournament were different for retailers who did not participate in the 2002 contest. Because all retailers had to participate in the second tournament, this evidence will provide a test for the self-selection hypothesis.

Results are summarized in Table 2. Columns 1–4 pool all 2002 observations, including a set of monthly time dummies to control for changes in market conditions. Columns 5–8 include only observations up to the first month of the tournament (January–May 2002). Finally, columns 9–11 compare the change in behavior between December 2002 and January 2003 of participants and nonparticipants in the 2002 program.

The incentive effect of the tournament is captured by the coefficient on the *Diff-in-Diffs* variable, which takes the value of 1 whenever the individual is in the tournament as of the beginning of that month and 0 otherwise. The positive and significant coefficient in columns 1 and 5 suggests that retailers increase their effort in response to participation in the tournament. Furthermore, this effect is sizeable, as it represents an increase in the performance measure of 24.5% for participants during the tournament.[27] Moreover, in column 1 we include a variable (*Drop*) to control for the effort of participants who were disqualified from the program in the past. The fact that the coefficient on the *Drop* variable is small and insignificant suggests that retailers excluded from the tournament go back to effort levels that are similar to their precontest performance, as theory suggests and as is consistent with the evidence in Knoeber and Thurman (1994).

The last three columns in Table 2 present further evidence that the incentive effect observed is not due to self-selection. We compare the incentive effects of the 2003 contest for the retailers who did versus did not participate in the 2002 contest. If participating retailers in the 2002 contest self-selected into it based on higher expectations and a greater ability to win, we would expect them to show a higher increase in sales during the first month of 2003. The coefficient on *Diff-in-Diffs* (the interaction between the dummy for January 2003 and the dummy for being a contestant in 2002) is small and not significant. This result persists even when we include only participants who remained in the contest in December 2002 or only those who were dropped before that date, suggesting that nonparticipants in 2002 react to the 2003 tournament similarly to all other contestants.

In columns 2 and 6 we expand the previous difference-in-differences analysis to consider the impact of tournament size on incentives. To do this, we add a variable for the interaction between the *Diff-in-Diffs* variable and the size of the tournament as measured by the number of prizes. The coefficient on this variable reveals how the incentives of the tournament change when we increase the size of the contest while keeping the proportion of winners (approximately) constant at 10% of the initial number of contestants. Although limited to only eight different tournaments, the results suggest that the larger the number of retailers, the less successful the Reward

[27] This number is calculated from Table 2 as the total effect obtained in column 1 divided by the average output for participants prior to the tournament.

Table 2 Incentive Effect of Tournaments

	Point Sales	*Point Sales*	*High-End Products*	*Unit Sales*	*Point Sales*	*Point Sales*	*High-End Products*	*Unit Sales*	*Point Sales*	*Point Sales*	*Point Sales*
	1	2	3	4	5	6	7	8	9	10	11
Diff-in-Diffs	167.822	285.885	0.015	66.630	179.451	410.740	0.008	82.102	21.920	9.384	74.406
	(33.301)‡	(151.946)*	(0.008)*	(23.099)‡	(48.035)‡	(103.986)‡	(0.014)	(28.383)‡	(84.675)	(75.767)	(250.407)
Drop	10.986	12.142	0.003	−8.962							
	(31.480)	(85.663)	(0.008)	(21.755)							
Diff-in-Diffs × *Contest Size*		−13.958				−28.083					
		(7.368)*				(8.479)‡					
Participant									415.453	254.682	1,088.57
									(44.235)‡	(39.875)‡	(128.169)‡
January 2003									116.591	116.591	116.591
									(44.258)‡	(44.266)‡	(44.321)‡
Constant	535.641	535.641	0.046	392.967	535.641	535.641	0.049	392.967	205.671	205.671	205.671
	(29.045)‡	(14.149)‡	(0.004)‡	(16.468)‡	(24.298)‡	(24.236)‡	(0.004)‡	(13.459)‡	(16.267)‡	(16.270)‡	(16.290)‡
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Observations	9,228	9,228	7,437	9,228	3,845	3,845	3,065	3,845	1,484	1,302	722
R-squared	0.71	0.71	0.59	0.75	0.83	0.83	0.73	0.89	0.05	0.04	0.22

Notes. This table shows the effect of the introduction of the tournament on the effort level of the retailers in 2002. Dependent variables are *Point Sales*, *High-End Products* (the fraction of sales corresponding to the premium and specialty products), and *Unit Sales* (which is measured in units rather than points). An observation corresponds to an individual during a particular month. The *Participant* dummy takes a value of 1 whenever the individual participated in the tournament in 2002. The *Tournament Month* equals 1 if the tournament was in place at that particular month (May to December of 2002 for columns 1–8 and January of 2003 for columns 9–11). The *Diff-in-Diffs* dummy takes a value of 1 whenever the individual is in the tournament as of the beginning of that month. The *Drop* dummy takes a value of 1 if the retailer participated in the tournament but did not satisfy the qualification requirements in any month prior to the current period. Finally, the *Contest Size* corresponds to the number of prizes. Columns 1–4 use the full sample for 2002, and columns 5–8 use the first five months of that year (which include the first tournament month). Columns 9–11 use only December 2002 and January 2003 observations. The coefficients in these last three columns show the difference-in-differences estimate of the effect of the 2003 tournament on the retailers that had participated in 2002 vis-à-vis those who had not. It is a test of self-selection based on ex ante beliefs of chances to win. To control for the incentive effects of the last month of the tournament, we use three definitions of the participant group. In column 9 all 2002 participants are included. Column 10 includes only those 2002 participants who dropped before December. Finally, column 11 includes only those 2002 participants who remained in the tournament in December 2002. Robust standard errors are in parentheses (clustered by tournament in columns 2 and 6).

*, †, ‡Significant at the 10%, 5%, and 1% levels, respectively.

Program was in inducing greater effort.[28] This is consistent with the results obtained by Gibbs (1996). He shows that when the distribution of the shocks is symmetric and unimodal, incentives decrease as the number of participants increases so long as the fraction of prizes is small and kept constant (Proposition 6). The suggestion that incentives are lower in larger tournaments is borne out in our setting, in which roughly one prize was awarded for every ten participants.

Finally, in columns 3, 4, 7, and 8, we focus on the multitasking aspect of the incentive program. The tournament encouraged retailers to promote high-end products through the point system described earlier, in which high-end products received a higher weight in the performance calculations. We therefore want to understand to what extent the increase in performance is attributable to an effort to reach more clients as opposed to an effort to upsell the clients reached. To measure the extent of upselling, we look at the fraction of sales in high-end products for each retailer in each month (columns 3 and 7). We measure effort to reach more clients as the total monthly sales in units of product rather than points (columns 4 and 8).[29] The positive coefficient on *Diff-in-Diffs* in column 3 indicates that the tournament made participants shift their effort toward high-end products, although the effect is not significant for the first month (see column 7). Moreover, its economic magnitude (increasing the share of premium products from 4.5% to 6.0%) is quite sizeable. On the other hand, columns 4 and 8 show a positive and significant effect of the contest on total unit sales, suggesting this may be the main driver of the tournament-induced performance increase. The asymmetric results for the two dimensions of effort are consistent with the idea

[28] We also tested whether this result was due to systematic differences in the distribution of contestant size in each region or in the way performance targets were set. In regressions not tabulated here, we included the monthly target as an explanatory variable to proxy for these effects, but it did not change the results reported.

[29] We do not directly observe the number of clients reached, but the number of units sold per client is a derivative of demand over which retailers have little influence.

that retailers find it more difficult to influence consumers to upgrade their purchase than to reach additional consumers. The business-to-business nature of our setting, with well-informed end consumers, may contribute to our results. However, these results raise questions about the costs to the customer-facing employees of widespread initiatives to increase cross-selling and upselling. Future research seeking clarification of these issues should probably be conducted in a consumer retail setting.

In summary, the empirical evidence suggests that the contest increases the effort of participating retailers as long as they remain in the tournament. Further evidence shows that this incentive effect decreases in intensity as the size of the tournament increases. Finally, retailers respond to the multitasking dimension of the program by increasing their efforts to reach more clients and to a lesser extent by upselling those clients that they reach.

5.2. Dynamic Incentives and Rank Distance

In this section, we analyze how the retailers' tournament rankings affect the level of effort they exert. We showed in §3 that retailers' motivation (and hence effort) should have a ∩-shape. Thus we should expect that the higher leading contestants climb, the less effort they will exert. We should also expect that the further contestants fall behind, the less effort they will exert. Motivation should be highest for retailers in intermediate positions. We also test the multitasking dimension of the dynamic incentives, i.e., how retailers distribute their efforts between reaching more clients and upselling clients depending on their interim ranking.

Whereas in the previous section we pooled all the observations of participating and nonparticipating retailers, in this section we will limit our analysis to those retailers who participated and remained in the contest for any given month. Table 3 summarizes the results of our tests. The first column shows the OLS estimates on the differenced equation, which though biased are provided for comparison.[30] The remaining columns provide the Arellano and Bond (1991) estimates for the dynamic panel data model, with column 2, also for comparison purposes, treating the distance measure as exogenous. All other columns treat this measure as predetermined, using its lags as instruments.

The empirical evidence shows that a leading retailer's distance from losing positions has a negative impact on its effort, suggesting that participants lose motivation as their chances of winning increase, as our theory predicts. Conversely, distance to a winning position for a trailing retailer has a generally positive and significant effect on the contestant's effort. Thus, we do not observe the expected demotivation effect in the set of losing retailers.

To see whether retailers in nonwinning rankings exhibited demotivation at any point, we added a quadratic term to *distance for trailers* in column 4. The coefficient is negative and significant, suggesting that this increase in motivation is indeed smaller as we move down the rankings.[31] However, the coefficient is also very small and suggests that the maximum level of effort should occur at a distance of around 70 places to the winners. Most likely, this low coefficient is due to the attrition in the sample that prevents us from properly observing and identifying any demotivation effect. Indeed, there are very few observations with a large trailing distance. For this reason, and because *distance for trailers* seems to have a positive effect for the range we observe, we keep the specification without the quadratic effect from now on. We defer to future empirical research the in-depth analysis of the demotivating effects of low rankings, which would need to be addressed in a setting with higher observed variability in trailing distances.

We conjecture that two factors may be contributing to the absence of demotivation in retailers with smaller trailing distances. First, the reinforcement theory of learning (see Skinner 1969) suggests that retailers will improve their performance after receiving a low ranking, which they interpret as a signal that they are not being as effective as other retailers. Second, if trailing retailers believe their chances of winning are higher than their ranking suggests, both the expectancy-valence theory (Vroom 1964) and the goal-setting theory (Locke and Latham 1990) would predict them to exert a greater effort than they would if their expectations were in line with their ranking.[32] We leave for future work the exploration of the theoretical underpinnings of this effect.[33]

[30] Because there is a positive correlation between a participant's ranking and past output (and hence shock), we should expect the OLS coefficient to be biased upward for *distance for trailers* and downward for *distance for winners*.

[31] We also added a quadratic term for *distance for winners*, but it was not significantly different from zero.

[32] We thank an anonymous referee for suggesting a mechanism by which these beliefs can be formed: Trailing retailers may attribute the performance of leading retailers to overconfidence that leads them to buy more products than are justified by their real sales potential. In consequence, trailing retailers may expect leading retailers to underperform in the next period, and they will be obligated to liquidate excess inventory and subsequently reduce their purchases. Based on this conjecture, trailing retailers may expect a higher likelihood of succeeding in the contest and therefore may be motivated to exert more effort than seems commensurate with their ranking. Note that for this mechanism to work, it is not necessary for leading retailers to actually drop in performance, only for trailing retailers to believe that they will.

[33] For instance, we could incorporate retailers' biased beliefs about their ability in the model in §3.3. Such beliefs could capture

Table 3 Tournament Design and Dynamic Incentives

	Point Sales	*Point Sales*	*Point Sales*	*Point Sales*	*Point Sales*	*Point Sales*	*Sales over Target*	*High-End Products*	*Unit Sales*
	1	2	3	4	5	6	7	8	9
Lagged Output		0.169	0.151	0.125	0.041	0.110	−0.009	0.082	0.053
		(0.159)	(0.150)	(0.159)	(0.117)	(0.139)	(0.054)	(0.065)	(0.062)
Leading Distance	−66.350	−83.799	−54.980	−42.013	−58.293	−43.848	−0.074	−0.006	−19.722
	(12.826)‡	(18.043)‡	(22.254)†	(21.916)*	(25.767)†	(20.659)†	(0.049)	(0.003)†	(10.898)*
Trailing Distance	7.607	11.310	4.095	10.565	6.262	3.838	0.018	0.0003	0.627
	(1.623)‡	(3.345)‡	(2.255)*	(4.976)†	(2.830)†	(2.155)*	(0.004)‡	(0.0002)	(0.691)
Trailing Distance squared				−0.072					
				(0.039)*					
Target						0.142			
						(0.028)‡			
Constant	−11.860	9.694	5.591	10.797	29.099	7.417	0.088	0.001	−16.842
	(16.274)	(27.501)	(14.069)	(12.898)	(14.697)†	(13.763)	(0.031)‡	(0.001)	(3.410)‡
Observations	3,183	3,183	3,183	3,183	2,452	3,183	3,165	3,169	3,183
Number of individuals	496	496	496	496	405	496	491	493	496
AR (1)		−4.27‡	−4.44‡	−4.34‡	−3.79‡	−4.55‡	−3.52‡	−6.98‡	−3.70‡
AR (2)		−0.63	−0.66	−0.69	−0.71	−0.73	−1.24	−0.44	−1.44
Sargan test (*p*-values)		0.000	0.682	1.000	0.586	0.690	0.164	0.993	0.696

Notes. This table shows the dynamic effects of the tournament. Dependent variables are *Point Sales* and *Sales over Target* (both measured in points), *High-End Products* (the fraction of sales corresponding to the premium and specialty products), and *Unit Sales* (which is measured in units, rather than points). An observation corresponds to an individual during a particular month. *Leading Distance* is the difference (positive) between the ranking of a player that was among the set of leaders in a particular month and the lowest ranking that would receive a prize, and zero for retailers in trailing rankings. *Trailing Distance* is the same difference (positive) for a player that had a rank lower than the number of prizes, and zero for retailers in leading rankings. Column 1 shows the OLS estimates on the differenced equation. Columns 2 onward report the Arellano and Bond (1991) estimates for the dynamic panel data model. Column 2 treats the distance measures as exogenous, but all other columns treat them as predetermined and instrument them with their lags. With regard to the data used, columns 1–4 and 6–9 use the full sample of participants, whereas column 5 drops the first three months of each tournament. Robust standard errors are in parentheses.

*, †, ‡Significant at the 10%, 5%, and 1% levels, respectively.

We want to be sure that the results we find are caused by the dynamic nature of the contests and are not due to a mis-specified test or to other factors. To validate our econometric specification, we perform two tests. First, the table shows the presence of a significant negative first-order autocorrelation in the differenced errors but no second-order autocorrelation, as expected if the errors are i.i.d., as assumed. Second, the Sargan test of overidentifying restrictions also has insignificant *p*-values (except for the model treating the distance measures as exogenous in column 2) and hence cannot reject the hypothesis of valid instruments. Both of these tests suggest the econometric model in column 3 is then most appropriate for analyzing the data. Moreover, the absence of second-order autocorrelation of errors, together with the insignificant coefficient on the lagged output, rule out reversion to the mean as an explanation of our results.

In columns 5–7 we perform several robustness tests. First, we consider the possible effects of attrition by running the same regressions without the first three months of each edition of the tournament, when most of the attrition takes place. The results in column 5 show that nothing changes after these periods. Columns 6 and 7 check the robustness of the results when allowing for target effects. First, we introduce the target as an explanatory variable and find no appreciable difference to the distance coefficients.[34] And second, we run the same original model but use sales over target as our measure of output. Again the same pattern emerges, although the coefficient on the leading distance becomes marginally significant. Consistent with the theory, these regressions suggest that the qualification requirements cannot explain our results. The participants remaining after the first three months should be less pressed by these requirements because fewer of them drop afterward. Also, the results are robust to controlling for targets, which affect the severity of the restriction. Hence, the distance effects must be due to the dynamic nature of the contest. In analyses not reported we also tested for differences between large and small tournaments by interacting the distance variables with the size of the tournament. Although the coefficient on *distance for trailers* is still positive and significant, and the effect of *distance for winners* is

overconfident participants, and the new model could help determine how much overconfidence is needed to explain the observed patterns of behavior.

[34] Treating the targets as exogenous or predetermined makes no difference to the estimates here.

negative and marginally significant, none of the interactions are statistically reliable.

The last two columns explore the multitasking aspects of the sales effort—client reach and client upselling. The results suggest that both dimensions are present in the retailer's response to its ranking in the tournament as it changes its effort to reach more clients and to sell a higher proportion of premium products to the clients it reaches.

In summary, it seems that as a retailer climbs in the tournament rankings, winning the prize becomes increasingly likely and the retailer may thus lose motivation and reduce effort. Conversely, a retailer in a trailing position receives the signal that it is not being as effective as other retailers. It reacts by trying to catch up and only loses motivation when the gap to the winning position is sufficiently large. This result suggests that the contestants exerting the maximum effort are lagging behind those that are just on the edge of winning positions. Moreover, from a multitasking perspective, the change in effort induced by the ranking is channeled through an effort to reach more clients as well as an effort to upsell these clients. These results highlight the importance of a feature of this relative performance compensation system, the nonlinear structure of its rewards, which can cause differential responses to the contest over time.

5.3. Economics of the Tournament

The earlier evidence suggesting that the Reward Program was effective in inducing higher sales does not prove that its introduction was a financially sensible decision for the firm because profits were not taken into account. However, the fact that the firm continued the Reward Program after the period of study and continues to implement it to this date is evidence of management's belief in the value created by its incentive system—a value that consists in the margin generated by the additional units sold. Although the firm considered cost information, and consequently margin information, extremely confidential, it did indicate that the gross margin of all products exceeded \$25 per unit. Table 2 model 4 shows that for the 2002 contest, a retailer participating in the tournament sold on average 67 more units per month while in the tournament; whereas for the participants that were dropped from the tournament, sales fell by 9 units. From Table 1 panel B we can compute that of the 4,000 potential retailer-months in the 2002 tournament, just over 50% (2,033 retailer-months) corresponded to months in which participants were out of the tournament.[35] Using these numbers we can infer that the program generated over \$2.8 million in additional profits for the firm, clearly outweighing its costs in prizes (\$150,000) plus any reasonable estimate of its administrative costs.

[35] The 2,033 retailer-months for tournament drop is calculated as the sum of the number of previously participating retailers that failed to satisfy the qualification requirements prior to the start of each month (this number will be 205 for July, 341 for August, etc.).

6. Conclusion

This paper analyzes the incentive effects of several contests organized by a commodities company among its retailers. The contests—which ran over a one-year period with monthly ranking updates to the participants—awarded equal prizes to a fixed number of retailers who achieved the highest level of performance. The tournaments appeared to induce higher levels of effort among participants. However, these incentives were weaker for contests with a larger number of retailers. This result can explain some previous findings, such as the fact that the returns on becoming a CEO increase with the number of contenders for that position (e.g., Main et al. 1993). Our results suggest that to induce motivation, the size of the prizes must increase in large tournaments as compared to small ones. This holds even when the probability of winning a prize (as measured by the proportion of prizes per contestant) is kept constant, a stronger result than those found in previous studies.

As with any field study, the institutional environment of the research site limits the generalizability of our contributions. In particular, the absence of an independent provider of market information forces us to rely on the firm's own assessment of two important aspects of our study: first, that the introduction of the contest aimed to increase the power of the existing incentive system and was not a reaction to changes in the market; second, that the drop in sales volume experienced by the firm's retailers during the second year of the contest was caused by a market contraction, although the firm did manage to increase its share of industry sales, in part due to the contest. Finally, the limited window of the study—two years—leaves open the question of the results' sustainability.

We also offer some new insight into the costs of tournaments when they take place over several periods and involve dynamic considerations. We develop a simple model that shows how these tournaments make equal-ability contestants heterogeneous as the tournament progresses because past performance matters for the final outcome. This past performance, in turn, affects the level of effort participants choose to exert at any moment during the contest. Consistent with our predictions, we find that contestants in leading positions reduce their effort as their lead, and thus their confidence of winning, increases. On the other hand, we find that retailers in trailing positions exert additional effort trying to attain a winning rank, losing motivation only when the gap is very large.

Theory predicts that tournaments where participants differ in intrinsic ability produce lower effort levels, but this can be corrected with the implementation of a handicapping system. Our analysis shows that a similar bias is generated when a tournament takes place over several periods, because early performance gives some competitors a lead and places them in advantaged positions whereas others lag behind and find it more difficult to win. Hence at interim dates, tournaments become heterogeneous. This heterogeneity, however, is more difficult to correct from within the tournament framework. The managerial challenge is to preserve the motivation of all participants throughout the program, probably by finding a design that tries to preserve the linearity of incentives (as in Holmstrom and Milgrom 1987) or possibly by other means—altering the structure of prizes, changing the interim feedback given to participants, choosing the handicaps differently for winners and trailers, or adjusting the length of the tournament. Studying the effects of these features of dynamic tournament design is a promising avenue for future research.

Finally, we analyze the effects of the multitasking dimension of the contests, which allows retailers to improve performance by either selling more units or selling a higher proportion of premium products. We find that although retailers increase their effort in both dimensions, most of the performance improvement is achieved by increasing the number of clients reached (the number of units sold) rather than by upselling more to all clients. This is an important effect given the current strategic emphasis of many firms on upselling and cross-selling current customers. We leave it to future research to examine the contingencies that make it more attractive to change existing clients' behavior than to persuade new clients to buy the firm's product.

Acknowledgments

The authors are grateful to Shannon Anderson, George Baker, Steve Bond, Michael Gibbs, Chris Ittner, Thomas Kittsteiner, Paul Klemperer, Clare Leaver, Gábor Lugosi, Andreu Mas-Colell, Michal Matejka, Meg Meyer, V. G. Narayanan, and Lucy White for their comments. The authors also thank seminar participants at the American Economic Association 2005 meetings, Brunel University, Cornell University, Global Management Accounting Research Symposium 2004, Harvard Business School, IESE, London Business School, Management Accounting Section 2005, Massachusetts Institute of Technology, Queen Mary University of London, Real Colegio Complutense, Tinbergen Institute Rotterdam, Universidad de Alicante, Universitat Pompeu Fabra, University of California Los Angeles, University of Michigan, University of Oxford, and University of Southampton. The second author acknowledges the financial support of the Harvard Business School Division of Research, and the first author acknowledges the financial support of the Ministerio de Ciencia y Tecnología (SEJ2007-64340).

References

Al-Najjar, N. I. 2008. Large games and the law of large numbers. *Games Econom. Behav.* **64** 1–34.

Anderson, T. W., C. Hsiao. 1982. Formulation and estimation of dynamic models using panel data. *J. Econometrics* **18** 47–82.

Antle, R., A. Smith. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *J. Accounting Res.* **24**(1) 1–39.

Arellano, M., S. Bond. 1991. Some tests of specification for panel data: MonteCarlo evidence & application to employment equations. *Rev. Econom. Stud.* **58** 277–297.

Bandiera, O., I. Barankay, I. Rasul. 2005. Social preferences and the response to incentives: Evidence from personnel data. *Quart. J. Econom.* **120**(3) 917–962.

Brickley, J., J. Zimmerman. 2001. Changing incentives in a multitask environment: Evidence from a top-tier business school. *J. Corporate Finance* **7** 367–396.

Bull, C., A. Schotter, K. Weigelt. 1987. Tournaments and piece rates: An experimental study. *J. Political Econom.* **95**(1) 1–33.

Casas-Arce, P., F. A. Martínez-Jerez. 2009. The all-pay auction with handicaps. Working paper, Universitat Pompeu Fabra, Barcelona, Spain.

Ehrenberg, R., M. Bognanno. 1990. Do tournaments have incentive effects? *J. Political Econom.* **98**(6) 1307–1324.

Feltham, G., J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *Accounting Rev.* **69**(3) 429–453.

Gibbons, R., K. Murphy. 1990. Relative performance evaluation for chief executive officers. *Indust. Labor Relations Rev.* **43** 30–51.

Gibbs, M. 1996. Promotions and incentives. Working paper, University of Chicago, Chicago.

Hannan, L., R. Krishnan, A. Newman. 2008. The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *Accounting Rev.* **83**(4) 893–913.

Holmstrom, B. 1982. Moral hazard in teams. *Bell J. Econom.* **13** 324–340.

Holmstrom, B., P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* **55**(2) 303–328.

Holmstrom, B., P. Milgrom. 1991. Multi-task principal-agent analyses: Incentive contracts, asset ownership, and job design. *J. Law, Econom., Organ.* **7** 24–52.

Knoeber, C., W. Thurman. 1994. Testing the theory of tournaments: An empirical analysis of broiler production. *J. Labor Econom.* **12**(2) 155–179.

Lazear, E. 2000. Performance pay and productivity. *Amer. Econom. Rev.* **90** 1346–61.

Lazear, E., S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *J. Political Econom.* **89**(5) 841–864.

Locke, E. A., G. P. Latham. 1990. *A Theory of Goal Setting and Task Performance*. Prentice-Hall, Englewood Cliffs, NJ.

Main, B., C. O'Reilly, III, J. Wade. 1993. Top executive pay: Tournament or teamwork? *J. Labor Econom.* **11**(4) 606–628.

Matsumura, E., J. Shin. 2006. An empirical analysis of an incentive plan with relative performance measures: Evidence from a postal service. *Accounting Rev.* **81** 533–566.

Meyer, M. 1992. Biased contests and moral hazard: Implications for career profiles. *Annales d'Économie et de Statistique* **25/26** 165–187.

Moldovanu, B., A. Sela. 2001. The optimal allocation of prizes in contests. *Amer. Econom. Rev.* **91**(3) 542–558.

Müller, W., A. Schotter. 2009. Workaholics and dropouts in organizations. *J. Eur. Econom. Assoc.* Forthcoming.

Murphy, K. 1999. Executive compensation. O. Ashenfelter, D. Card, eds. *Handbook of Labor Economics*, Vol. 3B, Chap. 38. North Holland, Amsterdam, 2485–2563.

Nickell, S. 1981. Biases in dynamic models with fixed effects. *Econometrica* **49** 1417–1426.

Orrison, A., A. Schotter, K. Weigelt. 2004. Multiperson tournaments: An experimental examination. *Management Sci.* **50**(2) 268–279.

Rosen, S. 1986. Prizes and incentives in elimination tournaments. *Amer. Econom. Rev.* **76** 701–715.

Skinner, B. F. 1969. *Contingencies of Reinforcement*. Appleton-Century-Crofts, New York.

Slade, M. 1996. Multitask agency and contract choice: An empirical exploration. *Internat. Econom. Rev.* **37** 465–86.

Vroom, V. H. 1964. *Work and Motivation*. John Wiley & Sons, New York.