Incentives and rank effects in managerial tournaments

Julia Shvets*
University of Cambridge, Christ's College

December 2017

Abstract

When firms use relative performance pay in which they rank employees, an employee's behaviour may respond to the rank they get. What is the relative importance of rank effects compared to monetary incentives? What is the direction of rank effects? Arguably, a bad rank may generate desire to catch up, or it may discourage further effort.

In this paper we address these questions by analysing store managers in a large firm where the bonus is determined through a high powered tournament. We study managers' response to feedback about their rank. In this tournament, the bonus is a step function of rank, and so marginal incentives and rank have a non-monotonic relationship, allowing us to separate the impact of incentives from that of rank on behaviour of managers.

First, we find that managers ignore marginal incentives, but respond to rank. Second, their response suggests desire to catch up: when managers get a bad rank on either profit or service, they respond by improving performance. This response is monotonic in rank. Importantly, we show that managers achieve these improvements by making corresponding changes to labour and production, the key input variables directly under their control.

*js591@cam.ac.uk. I thank the Keynes Fund for financial support. I am grateful to Jordi Blanes i Vidal, Antonio Cabrales, Paweł Gola, Hamish Low, the participants of the Keynes Fund Research Day and the Judge Business School Economics and Policy seminar for their comments. Christian Holm provided excellent research assistance. I am grateful to the firm for sharing the data, and especially to Nick, Eliana, Stephen and Eric for all their help.

1 Introduction

1.1 Motivation

A growing body of literature from psychology and, more recently, economics argues that people care about their relative standing, or rank, among their peers.

This may have implications for tournament incentive schemes used by many organizations to reward performance; particularly as these schemes often provide regular feedback on rank to their employees. And yet, there is little relevant evidence: most of what we know about people's concerns regarding their rank comes from the lab or from settings that do not involve immediate monetary incentives. So it is an open question whether these rank concerns still manifest themselves when large financial rewards are at stake. For example, if monetary incentives and rank concerns call for different actions, will rank concerns dominate or will they take the back seat to incentives?

This unexplored question is the first of the two that we address in this paper. The second question we investigate is the direction of rank effects. Apriori, the impact of feedback about rank is not obvious: receiving a bad rank may either discourage effort or, conversely, generate desire to catch up. The few existing results from the lab and non-monetary tournaments have been somewhat contradictory, leaving open the question about what to expect in the workplace.

To address these questions, we anayze behaviour of store managers in a large food and beverage chain, where the managers face a high powered tournament incentive scheme. We exploit the rules of the tournament to separate the effects of incentives from those of rank. To preview our results, we find that *on the margin* managers generally do not respond to the monetary incentives of the tournament. However, they do respond to rank. More specifically, the managers who receive a worse rank on a particular performance measure make changes that lead to performance improvements on that measure.

So, in a high stakes real world tournament, we find that incentives take the back seat to rank concerns, on the margin allowed by our data. Furthermore, the managers' response to rank suggests that getting a bad rank leads to desire to catch up, rather than to discouragement.

1.2 What we do

The firm determines the bonus of their store managers using a quarterly tournament. The managers are given feedback on their performance two weeks before the end of the quarter, and this feedback creates variation in both marginal incentives and relative rank that the managers face as they go into the last two weeks of the tournament. We study the managers' response to this feedback, separating the impact of rank from that of incentives.

So far, to the best of our knowledge, the empirical literature on tournament incentive schemes has not been able to separate these two effects. To see why this is challenging, consider what the theory tells us about who will faces stronger incentives in a tournament. The standard result is that marginal incentives are strong for people who find themselves close to getting the prize, and weak for those who are further away, either because they lag too far behind or because they have a strong lead (Lazear and Rosen (1981), Casas-Arce and Asís Martínez-Jerez (2009), Ayoagi (2010), Ederer (2010)). Empirically, however, marginal incentives are often strongly correlated with rank – for example, in sports tournaments where only top people get prizes. This makes it difficult to empirically disentangle the effect of incentives from the effect of rank concerns.

We tackle this challenge by exploiting the rules of our tournament. In it, managers are ranked on three performance measures and the rankings are then divided into bands. Within a band, all managers are awarded the same contribution towards their bonus, and hence bonus is an (increasing) step function of performance. So here, theory implies that managers that who end up close to the boundary of the band on either side of it will face stronger incentives in the run up to the tournament. Hence, in our setting, the relationship between the strength of incentives and performance is non-monotonic, while rank always moves together with performance. This allows us to separate empirically the influences of incentives and rank on managers' actions.

To better understand our results, it is useful to give some background on what our managers do. In addition to managing labour (on average, 14 people in a shop at any one time), the manager decides how much labour to hire and how to divide it between two main activities: production and service. The firm evaluates its managers on three main measures: sales, profit and service. In the short run, it is harder for managers to influence sales than profit or service, both of which can be affected by changing total labour or its allocation.

We analyze how the three key performance measures – sales, profit and service – change in response to feedback. While average performance improves following feedback, there is also substantial heterogeneity. Taking care to address concerns about serial correlation in performance, we relate changes in performance in the last two weeks of the tournament to the contents of the feedback. Importantly, feedback contains information on rank and incentives (proximity to bonus band boundary) *for each performance measure.*

First, we show that managers who find themselves facing steeper monetary incentives on a given measure do not improve performance by more than the rest. So we find that on the margin, there is no evidence of managers responding to incentives.

Second, for profit and service, we find that managers respond to rank contained in their feedback. When service rank is worse, they improve service; when profit rank is worse, they improve profit. However, sales do not respond to rank consistent with sales being hard to change in the short run.

An important feature of our data is that, unlike in many empirical studies of tournaments, in addition to performance of managers, we observe key input decisions of our managers. More specifically, we have data on total labour, and output-labour ratio. Using a simple framework to capture how outcomes and managerial input decisions are related to each other, we show that in the short run, managers who want to improve profit should choose lower labour and higher output-labour ratio; while if they want to improve service they should do the opposite.

Using the data on observed input decisions, we verify that our performance results are indeed driven by the actions of managers. First, we confirm the absence of response to incentives in performance by showing that managers who face steeper incentives do not change inputs in expected ways. Second, we find that input decisions do repond to rank, primarily in profit and service. Finally, these responses are consistent with our performance results: when a manager ranks badly on profit, they cut labour and increase output-labour ratio, consistent with the desire to improve profit. Conversely, if a manager ranks badly on service, they hire more labour and let output-labour ratio fall. Hence, both performance and input decisions respond to rank, and they do so in ways that are consistent with each other.

To sum up, we find that managers do not respond to marginal incentives, but they respond to relative rank, and the direction of this effect suggests desire to catch up among those who lag behind, rather than discouragement. This is robust to inclusion of manager fixed effects and addressing concerns about potential serial correlation in performance, such as that due to working in spurts or to reversion to the mean.

1.3 Literature

1.3.1 Response to incentives in tournaments

The novelty of this paper is that it separates the effect of monetary incentives from those of relative rank by looking at how managers respond to feedback in a real world, high stakes tournament. In doing so, we contribute to the literature which studies response to marginal incentives in tournaments with feedback. Taken together, these papers provide some support for the theory of tournament incentives ('the theory' thereafter); however, they also show that often people behave in ways that are not consistent with responding to marginal incentives.

Delfgaauw et al (2014) conduct a field experiment in a retail chain where they create a quasi-random variation in the intermediate rank of competitors. They find that shops closer to the prize improve performance by more, consistent with the theory. Casas-Arce and Asís Martínez-Jerez (2009), who also study a retail tournament, find that, in line with theory, participants with a large winning lead reduce their efforts; however, contrary to the theory those who lag behind do not reduce their efforts, unless the lag is very large.

Ehrenberg and Bognano (1990) find surprizing results in golf tournament data: the players who face higher marginal incentives go on to score lower in later rounds. Genakos and Pagliero (2012), who study risk taking and performance in weightlifting competitions, show that despite being close to winning, weightlifters with interim ranking that is near to the top underperform their competitors with worse ranking. Again, this is contrary to what we would expect if people simply responded to marginal incentives.

In laboratory, Müller and Schotter (2010) vary ability to show that, as predicted by theory, competitors with a large lag drop out; however, competitors a large lead work 'too hard' relative to the predictions of the theory. Eriksson, Poulsen and Villeval (2009) explore several incentive schemes in the lab, and find that in tournaments reactions to feedback are inconsistent with response to marginal incentives: 'underdogs never quit, and front runners never slack off'.

Taken together, these findings suggest that in addition to creating certain marginal incentives, feedback about rank in tournaments may also have other effects. Given emerging evidence on rank concerns, some of which is discussed in the next section, it seems plausible that some of these effects may have something to do with people caring about their relative rank per se. By and large, the papers above cannot speak to this: in most of these settings, rank and incentive effects are hard to separate because the two things are highly correlated, either because prizes are only given to the few top performers (for example, in Casas-Arce and Asís Martínez-Jerez (2009)) or prizes are continuous and marginal incentives are monotonic in performance (for example, in Ehrenberg and Bognano (1990)). Our key contribution is to overcome this problem, thanks to the rules of our tournament which allow us to decouple rank and incentive effects.

Furthermore, in most of the papers above, response to feedback is inferred from change in performance, which is assumed to be at least in part due to change in effort. In contrast, we observe not only performance but also some key input decisions of the managers. By analyzing changes in these decisions along side performance, we are able to look at managers' response to feedback in a much more direct way.

Finally, unlike in most of the settings mentioned above, we observe multiple dimensions of managerial performance, which are differentially affected by managers' input decisions. This provides an important consistency check for our analysis. And so the fact that our findings are, indeed, in line with these differential predictions lends further confidence to our results.

1.3.2 Response to rank feedback

Another strand of literature which is closely related to our study, and from which we drew a lot of inspiration, investigates effects of relative performance feedback in settings where payoffs are not immediately related to rank. It supplies key evidence that rank per se may be important to people.

Like the papers discussed above, and the present paper, Gill et al (2016) analyze response to tournament feedback. Crucially, in their lab experiment the pay is unrelated to performance, allowing elicitation of pure preferences over rank. By randomly breaking ties in performance, the authors generate exogenous variation in feedback, to minimize problems due to serial correlation in performance. They find that post-feedback performance has a U shaped relationship with feedback: those who are in the top and in the bottom positions work harder, compared to the people in the middle, the pattern the authors dub 'first place loving and last place loathing'.

In addition to analyzing the content of feedback, whether people care about rank can be studied by looking at how their behaviour depends on whether they receive the feedback or not. There has been a number of such studies, with somewhat contradictory results. In the lab, Azmat and Iriberri (2015) find that introduction of relative performance feedback has no effect on performance under flat rate, but improves it under piece rate. Eriksson, Poulsen and Villeval (2009) show no impact on average under piece rate or tournament. In the workplace where rank is not tied to pay, Blanes i Vidal and Nossol (2011) find that introducing relative feedback improves performance, while Barankay (2012) seems to contradict this by showing that *removing* feedback does the same. In education, high school students in Azmat and Iriberri (2010) work harder after provision of feedback, but stop when feedback is removed. Azmat et al (2016) show that among university students, the reaction to introduction of feedback depends crucially on students' initial expectations of their relative performance. On average, feedback worsens exam results, but this effect is temporary.

Among other things, Azmat et al (2016) findings suggest that studies looking at introduction or removal of feedback policies, unless done over a long time horizon, may pick up short-lived surprise reactions. Our study complements this literature by supplying evidence on rank feedback effects in a steady-state environment, where most managers have extensive experience with the tournament and feedback.

More generally, alongside these papers, our study contributes to a growing body of evidence that people care about rank: we show that even in a high-stakes, workplace tournament where rank is directly linked to pay-off, rank feedback continues to influence behaviour over and above marginal incentives.

Along side estimating rank and incentive effects of a tournament, we also show that managerial performance improves across the board as the tournament day nears. In this way, our study is consistent with earlier findings that in dynamic settings, performance often improves towards the pay date (Oyer (1998), Kaur et al (2016)).

1.4 Plan for the rest of the paper

The rest of this paper is organized as follows. Section 2 describes the context of our study, using a simple framework to link performance outcomes and observable input decisions. It also describes the bonus scheme which form the basis for the empirical identification, laid out in section 3. Also in section 3, we state our empirical predictions and discusse the data. Our results are in section 4. Section 5 concludes.

2 Manager decisions and incentive scheme

Our aim is to better understand what drives managers by looking at how they respond to performance feedback in a quarterly tournament organized by the firm. The tournament, and hence the manager bonus, is based on three main performance measures: sales, profit and service. Two weeks before the end of the quarter, for each measure m, the managers get feedback on their relative position which determines both m-specific incentives and rank that a manager faces going into the last two weeks of the tournament. We then look at whether in these last two weeks managers promote performance on measures where they (1) face steeper incentives and/or (2) face low or high ranks in feedback.

We complement this performance analysis with the analysis of manager's input decisions in the last two weeks before the tournament. The main ways in which a manager can influence outcomes are hiring and allocating labour across tasks, and exerting effort. One advantage of our data is that we are observe some the key aspects of these decisions. In section 2.1, we describe the managers' job and their input decisions using a simple framework, and then show how we would expect the input variables *observable to us* to change depending on which performance measure the manager wants to promote. Then subsequently we will take these predictions to the data to see whether managers' response to incentives and rank can be traced not only to changes in performance but also to corresponding input decisions.

In section 2.2, we describe the firm's incentive system and feedback that managers receive, which sets the basis for our identification strategy (described in section 3).

2.1 Manager decisions and outcomes

2.1.1 Context in a simple framework

In each shop, the manager is in charge of a team of employees, who are engaged in two main tasks: production of output for sale (food and beverage) and customer service. We denote the two activities by P and S respectively, quantity of output produced by Q_P, and service produced by S. The key decision of the manager is how much labour L_i to employ for each of the two tasks where $i = \{P, S\}$. An average store has 14 workers at any given time, and 20-22 workers on its books, some of whom are part time. This excess capacity means that a manager can change labour quite quickly.

In addition to choosing L_S and L_P, the manager can exert effort to organize processes better and motivates employees, which can affect labour productivity in the two tasks.

To a good approximation, the managers do not make decisions related to raw materials. They also have no control over the prices of output, which are set centrally. The managers have a limited degree of control over the varieties of output they produce; we abstract from this for tractability.

We will now make assumptions about how the key outcome variables depend on labour and on each other. Let

$$S = g(L_S) \tag{1}$$

$$Q_P = f(L_P) \tag{2}$$

both of which are increasing functions and $f(.)$ is concave. This is plausible, since the store is subdivided into production and service areas which cannot be altered in the short to medium run.

The following describes the managers' sales:

$$Q_S = x[\min(\bar{Q}, Q_P) + \alpha(S)\max(Q_P - \bar{Q}, 0)] \tag{3}$$

where x is the price of output, and the term $\bar{Q}$ captures the possibility that a store may be demand constrained. The degree to which this constraint can be relaxed is given by $0 \leq \alpha(S) \leq 1$, allowing for α to be a function of service. We choose this functional form because many in the company believe that in the short run, the managers cannot influence demand, i.e. α is close to zero.[1]

[1]For example, when in a survey we asked the managers what was the easiest way to

Below, in section 2.1.2 we analyze manager choice for positive α, and then discuss the case when $\alpha = 0$.

Finally, profit is given by

$$\Pi = xQ_S - w(L_P + L_S) \tag{4}$$

where w is wage.

2.1.2 Analysis

The tree performance measures on which the manager is evaluated are given by equations (1), (3) and (4). Assuming they all enter the manager's objective with non-negative weights, and without specifying the objective function, we can conclude:

- For each $i = \{P, S\}$, all three performance measures weakly increase in L_i as long as $L_i < \text{argmax}_{L_i} \Pi$. Hence, the manager will set

$$L_i \geq \underset{L_i}{\text{argmax}}(\Pi) \text{ for both } i \tag{5}$$

- Once $L_i > \text{argmax}_{L_i}(\Pi)$, there are trade offs:
  - $\frac{\partial Q_P}{\partial L_P} \geq 0$ while $\frac{\partial \Pi}{\partial L_P} < 0$
  - $\frac{\partial S}{\partial L_S} > 0$ while $\frac{\partial \Pi}{\partial L_S} < 0$

So, when the manager wants to promote sales or service, L_P and L_S respectively will be higher and further away from the relevant $\text{argmax}\,\Pi$. If he wants to promote profit, L_P and L_S will be lower and closer to the relevant $\text{argmax}\,\Pi$.

If the manager wants to improve profit or sales he may also achieve this by increasing average productivity of L_P without changing labour. For example,

increase their bonus by a specified amount, 30% said through service, 17% said through profit, and only 10% mentioned sales (the rest answered this open-ended question without mentioning a specific performance measure).

if the manager re-allocates labour in production towards peak hours where the value of marginal product is higher, production per worker can increase without a change in labour.[2] Such changes are likely to require costly effort from the manager. Hence, changing average product of L_P is a key way in which the manager's effort can affect the outcomes.

2.1.3 Implications for observables

We have the data on the amount spent on total labour by each manager, $L \equiv w(L_P + L_S)$, but not on the two tasks separately. We also observe the value of production and so can calculate the value of production per dollar spent on labour, $q \equiv xQ_P/L$, to which we loosely refer as output-labour ratio. q can change for two reasons. First, it falls in L: by definition, $\frac{\partial q}{\partial L_S} < 0$ and, by concavity, $\frac{\partial q}{\partial L_P} < 0$. Second, for a given L it increases with improvements in average productivity of L_P (which, in turn, are driven by manager's effort).

Based on the trade offs we have identified in section 2.1.2 above, we can draw out the following implications for the observables L and q:

- If a manager wants to promote sales or service, he will choose a higher L; if he wants to promote profit he will choose a lower L.

- If a manager wants to promote profit, he will set a higher q (either through lower L or higher average productivity of L_P); if he wants to promote service, he will set a lower q (through increases in L); if he wants to promote sales, this may either imply a lower q (if L expands) or a higher q (if average productivity of L_P rises).

- If a manager is sales constrained (i.e. in equation (3) $\alpha = 0$ and $f(\text{argmax}_{L_P} \Pi) \geq \bar{Q}$), then the above conclusions still hold for profit and service. However, now sales cannot be changed using the inputs available to the manager.

[2] I am grateful to Michael Pollitt for this example.

2.2 Incentive scheme

The managers get paid a base salary but also substantial performance bonuses, awarded through a quarterly tournament. All managers get a bonus, and its size depends on how they fare in the tournament. The average bonus is roughly 20% of the base salary, while the top bonus is about 150% of base salary, making this a relatively high powered incentive scheme. Thus, managers face substantial incentives to understand the scheme and make decisions that lead to higher bonus.

The bonus is calculated by aggregating four store performance measures: sales growth, profit relative to a store specific target, customer service and an evaluation of store operations by an area manager. The service is measured using mystery shoppers which fill out a score card. The tournament is held at the end of each quarter based on the managers' cumulative performance for that quarter.

Indexing these performance measures by m, the bonus is calculated like this:

1. For each performance measure m, managers' outcomes y_{im} are ranked, and each manager i gets a rank r_{im}.

2. Each of the four rankings is divided into bands. Within each band b, all managers get assigned the same score s_{imb}. The score is higher for higher bands (near the top of the ranking).

3. For each i, bonus rate B_i is determined by

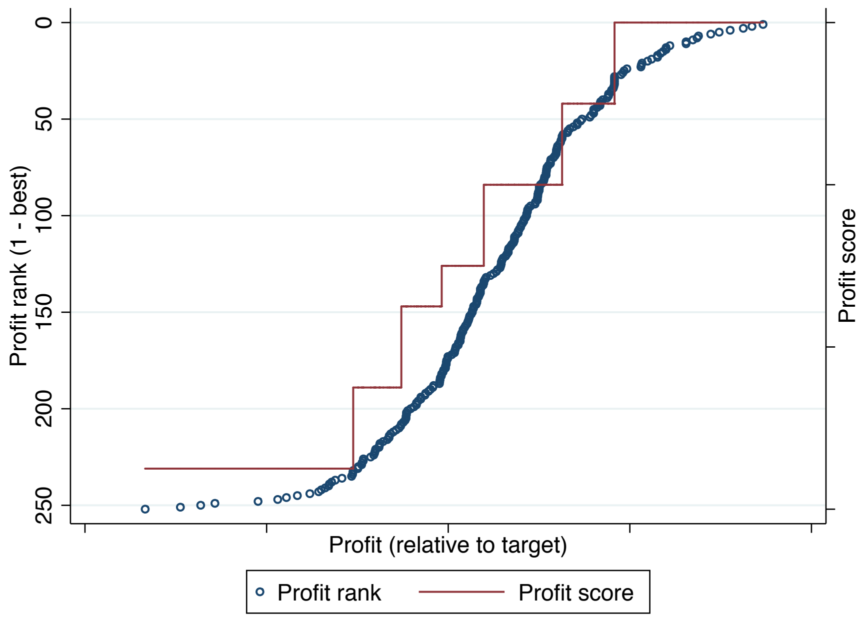
   $$B_i = \prod_{m=1}^{4} s_{imb}$$

   Also, each managers receives an overall ranking R_i =rank(B_i).

4. Rate B_i is normalized and applied to manager's salary to find the level of the bonus. A few managers with the lowest R_i (highest B_i) get additional bonus.

These rules imply that, for each m, the score is a step function of the performance on that measure. Therefore, for given performance on other measures, bonus is a step function of performance on measure m. Figure 1 shows an example of this using profit data from Q2 of 2014[3].

Figure 1: Profit rank and score



As we can see from the graph, while profit score and hence the manager's bonus jumps at certain levels of profit, the manager's rank changes smoothly and monotonically with profit. This is the key feature that allows us to to separate the effects of incentives from those of rank on manager behaviour: while rank always rises, incentives will be stronger nearer to bonus jumps, and so the strength of incentives does not always move in the same direction as the rank. We discuss this further in our empirical identification section (section 3).

We make two practical decisions with respect to the four performance measures included in the incentive scheme. First, recall that the firm measures

[3]The rank axis in figure 1 is reversed so that the best (lowest) ranks are at the top

sales and profit against benchmarks (sales growth relative to the same period last year and profit in excess of a shop-specific target, which changes very infrequently). From the point of view of our estimations, both past sales and profit targets are fixed, and outside the manager's control in the last two weeks of the quarter. What the manager can affect are current sales and profit, and so these are the variables we will focus on. Second, evaluation by an area manager takes place only once a quarter, which means there is no feedback and we cannot study it. So, henceforth, our study focuses on sales, profit and service outcomes.

2.2.1 Score bands

Typically, the ranking is divided into seven bands. The bands are drawn in such a way that top (highest) two bands have 10% of shops in them, the next two bands have 30% and 20% respectively, and the bottom three bands have 10% of shops in each[4]. An average band has 14 shops in it.

Marginal benefit that a manager i faces from jumping up one band on measure $p = j$ is given by

$$MB_{ijb} = \prod s_{bi,m \neq j} \Delta s_{ijb} \tag{6}$$

where Δs_{ijb} is the jump in score that manager i faces on measure j. Depending on quarter and band, the jumps Δs_{ijb} can take on four values in our data: 5, 10, 15 and 20 points on the score scale, or 1.5-6 percentage points of the final bonus rate. The mode and median jump is 10 points or 3 percentage points of the bonus rate. Since the average bonus rate is 20% of base salary, jumping up one band on one measure gives a sizeable improvement in pay. Figure 1 illustrates a typical configuration of jumps: 10 points for all bands, except for one band in the middle, where the jump is 5 points.

[4]In practice, the bands can be slightly narrower or wider, to ensure that shops with the same absolute performance fall into the same band.

2.2.2 Information and feedback

When managers join the firm, the company explains the rules of the tournament and has a reference document detailing these explanations. There is also full transparency about the outcomes: after the quarter end, managers receive information about their bonus and also a table listing the results for all managers in the tournament.

Weekly, the firm provides detailed feedback to managers on their performance. With thirteen weeks in a typical quarter, the firm provides feedback at the end of each of the first eleven weeks. In the first instance, managers get this weekly feedback electronically, and then they often discuss it with their area manager. No feedback is provided in weeks 12 and 13, and then, after the end of the quarter, the results of the quarterly tournament are announced in a meeting of all the managers participating in the tournament.

Weekly feedback provides the results of a hypothetical tournament based on managers' performance upto and including the current week. Each manager gets five tables, one for each measure (sales, profit, service, and area manager evaluation) and a summary table[5].

In each table covering a particular measure, all stores are listed in the order of rank (r_{im}) on this measure, as in the example for service in table 1. Absolute performance (y_{im}) and score (s_{imb}) are also reported, making it easy for the store manager to see whether they are close or not to the boundary of the bonus band, i.e. the point where their score jumps on a each measure. In our example in table 1, $x > z$ and the manager of store G is immediately below the bonus band boundary, while the manager of store F is immediately above it.

2.2.3 Regional and national tournaments

The firm runs two kinds of tournaments – one is conducted nationally, with all the shops in the country competing, and for the other, the country is

[5]Since area manager evaluation only occurs once a quarter, for most weeks and shops this feedback table is empty. For the purposes of calculating weekly overall bonus rate, the firm assumes a baseline score on this measure for all managers that have not yet been evaluated.

Table 1: Example of feedback table for service

Store	Service y_i	Service rank r_i	Score s_{ib}
.	.	.	.
.	.	.	.
.	.	.	.
D	88	10	x
E	86	11	x
F	85	12	x
G	84	13	z
H	82	14	z
I	81	15	z
J	80	16	z
.	.	.	.
.	.	.	.
.	.	.	.

split into 2-4 regions, with a separate tournament in each. Until Q4 of 2011, all tournaments were regional. Thereafter, quarterly tournaments alternate between regional ones (in odd quarters) and national ones (in even quarters). The rules of regional and national tournaments are nearly always the same. In our data 4 out of 18 quarterly tournaments are national, and the rest are regional[6].

The weekly feedback system was put in place when all tournaments were regional, and so the rankings and scores that it provides are based on the assumption of a regional tournament. This means that while absolute performance feedback is equally relevant in all quarters, relative performance measures (ranking and score) are a more accurate forecast of the quarterly outcomes when the actual tournament is regional. Empirically, our data are dominated by regional tournaments where the weekly feedback is the perfect match for the actual tournament, and our key results are unchanged if we drop the national tournaments from our sample[7].

[6]This in part because we exclude all Q4 tournaments from our data - see section 3.2.

[7]Also, the managers have access to weekly feedback for all regions, and so arguably they can compute their 'correct' relative performance feedback even for the national tournament.

3 Empirical identification and data

3.1 Empirical identification

In week 11, managers receive information about their relative rank on each of the three performance measures, sales, profit and service. In addition, because they see the entire ranking which shows where the bonus bands had been drawn, they also know how close they are to the boundary of the band, and this determines their incentives.

We want to identify the effects of incentives and rank received in week 11 on behaviour of managers in weeks 12 and 13. In this section, we first expand on how rank and incentive effects can be separated from each other in our setting (section 3.1.2). Second, we discuss how we deal with the challenges posed by potential overtime correlation in performance (section 3.1.4).

3.1.1 Separating rank and incentives

The key insight from the theoretical literature on tournaments is that incentives are strongest for competitors who find themselves, for one reason or another, nearer the prize. Conversely, those who have a strong lead or substantially lag behind, will face weaker incentives. This is because the former know their lead cannot be easily eroded, while the latter know their gap cannot be easily bridged.

Lazear and Rosen (1981) were first to show this in a contest with heterogeneous abilities. Subsequently, Casas-Arce and Asís Martínez-Jerez (2009), Aoyagi (2010) and Ederer (2010) extend this result to tournaments with intermediate feedback, where the past feedback, rather than ability, is the source of heterogeneity[8].

Recall that in our setting, the bonus is a step function of performance, because ranking is divided into bands. So, the implication of the above theoretical result for our prize structure is the strongest incentives will be experienced

[8] This result requires the assumption that performance shocks are iid and their joint distribution is symmetric and non-uniform.

by the managers who in week 11 find themsleves just below or just above a boundary of a bonus band. The former only needs a small effort to bridge the gap, and the latter might drop in bonus as a result of only a small shock.

This means that in our setting the strength of incentives oscillates with performance, which is what allows us to separate the influences of incentives and rank on managerial behaviour. Unlike, for example, in contests which have prizes only at the top, in our tournament incentives and rank do not always move together: while rank follows performance, incentives go up and down with it, depending on the distance from the boundary of the bonus band.

3.1.2 Regression equation and hypotheses

To capture that incentives are strongest near the boundary of the bonus band, we define a variable $I_{itm} = 1$ if week 11 feedback in quarter t puts manager i within two ranks on either side of this boundary on performance measure $m = \{Q, \Pi, S\}$, where Q is sales, Π is profit and S is service. In the results section, we will show that our findings are not affected by changes to the width of this window.

The rank for manager i in quarter t, r_{itm}, is defined in relative terms: the raw rank from week 11 feedback table for measure m is divided by the total number of stores in that table, to remove fluctuations in rank due to changes in the number of competitors. Rank is inversely related to performance: best performer has relative rank 0, worst – relative rank 1.

Finally, we define our dependent variable as the change in manager i's outcomes and input decisions in the last two weeks of quarter t as

$$\Delta y_{itj} = \frac{\sum_{w=12,13} y_{itjw}}{2} - \frac{\sum_{w=1}^{11} y_{itjw}}{11} \tag{7}$$

where $w = 1, .., 13$ is a week, and $j = \{Q, \Pi, S, L, q\}$ is the outcome or input decision of interest. Recall that L is total labour and q is output-labour ratio.

We estimate the following equations, one for each of the five js:

$$\Delta y_{itj} = \sum_m \beta_{mj} I_{itp} + \sum_p \gamma_{mj} r_{itm} + \epsilon_{itj} \tag{8}$$

Our hypotheses on the sign of incentive coefficients are these:

- For the three outcome variables $j = \{Q, \Pi, S\}$, we expect 'own' incentive coefficients, $\beta_{m=j,j}$, to be positive: higher incentives improve performance. Sales are harder to influence in the short run than the other two measures, and so we expect weaker effects on them.

- For input decisions L and q, the expected effects come from the framework in section 2.1:

  - Profit incentives lower L and increase q
  - Service incentives increase L and lower q
  - Sales are harder to change, so sales incentives may have no effect; if they do we expect them to increase L, while the effect on q is ambiguous.

It is harder to formulate hypotheses about rank coefficients, because we do not have a model of how people may respond to rank. Intuitively, we distinguish between two main possibilities:

1. *Desire to catch up.* Managers who get worse (higher) rank on measure m try to improve it, so for the outcome variables $j = \{Q, \Pi, S\}$, 'own' coefficients on rank should be positive, i.e. $\gamma_{m=j,j} > 0$. The expected impacts of ranks on input decisions should be the same as those of incentives above.

2. *Discouragement.* Managers who get worse (higher) rank on measure m get discouraged and give up trying to improve it, but do try improve when they get a good (lower) rank. This suggests $\gamma_{m=j,j} < 0$ for 'own' effects on outcomes. The expected impacts of ranks on input decisions should be the opposite of those of incentives above[9].

[9]Note that the first hypothesis is equivalent to saying that the marginal utility from rank

Table 2: Probability of crossing into the next band

	Sales	Profit	Service
Near band border	0.67	0.60	0.66
	(0.01)	(0.01)	(0.01)
The rest	0.63	0.48	0.43
	(0.01)	(0.01)	(0.01)
N	3,632	3,656	3,680

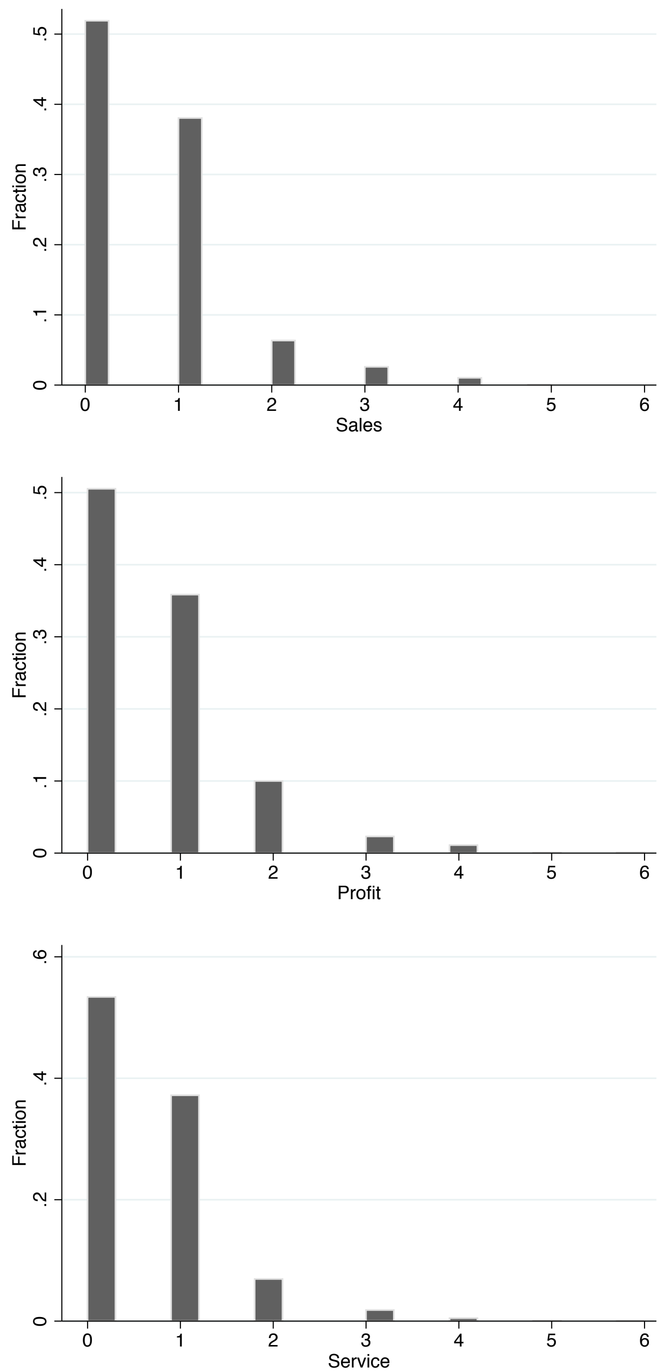
3.1.3 Checking assumptions

Implicit in our empirical strategy is the idea that in week 11, the relevant prize for the manager is the one closest to him (rather than the prizes that are several bands away). This is true if it is hard for the manager to move performance in the last two weeks of the quarter by more than one band. Figure 2 confirms that this is the case: for each of the three performance measures, 90% of managers either remain in the same band or move by only one band in the last two weeks.

When we say that the incentives will be stronger nearer the band border, our conjecture is that it is easier for people nearer the band border to cross it than for people who are further away. Indeed, this is what we see in table 2: for all three measures, the probability of the managers crossing into the next band is statistically significantly higher if you are near the border. For sales, the difference is 4 ppt and is significant at 5%, while for profit and service the difference is considerably bigger: 12-13 ppt, significant at 1% level. The fact that proximity to the border in sales has a lower impact is consistent with the idea that sales are harder for managers to influence compared to profit and service.

improvements falls towards the top of ranking, while the second that it rises towards the top.

Figure 2: Number of bands travelled in last two weeks*



3.1.4 Serial correlation in performance

A key challenge in identifying the effects of performance-based rank at time t on subsequent performance in $t-1$ is potential serial correlation in performance across time. A number of distinct of mechanisms can lead to this, and so we use several approaches to deal with this issue.

Differenced dependent variable

Suppose performance over time has a significant fixed component, for example due to fixed manager or shop characteristics, creating positive correlation in performance over time. By using first differences as our dependent variable, we net out this fixed effect.

Manager fixed effects

Still, differencing may not fully remove the correlation due to manager or shop fixed effects. For example, higher ability managers can be quicker learners and so the level of performance, which captures ability, will be positively correlated with the change over time (Gill et al 2016). To tackle this, we use manager fixed effects, which will pick up such fixed characteristics like ability.

Controlling for level of performance

Even when we use manager fixed effects, some types of manager characteristics can still cause a problem for our estimations. For example, the manager may be increasing their effort towards the end of quarter due to discounting of the future; managers who discount more may have lower week 11 performance and a bigger increase in the last two weeks (for example, Gill et al (2016) refer to this as 'working in spurts'). If each manager always discounts future in the same way, then manager fixed effects are sufficient to deal with this issue. However, the same manager may shift more effort towards the end of the quarter in some quarters and not others (for example, when the weather early in the quarter is particularly good), and this will generate a negative correlation between the level of performance in week 11 and the subsequent change, even if we only use within manager variation.

We address this problem by controlling in the regression for the absolute level of managers performance at the end of week 11. Hence, the correlation between the level of performance and the change, which can come from working in spurts, or from other channels, such as reversion to the mean, will be taken out. The remaining variation in rank will be from performance of other contestants, and so any effects of rank on subsequent performance we identify are the effects of the manager's relative standing among his peers, controlling for the level of his performance.

Disaggregated data

Our dataset has two key advantages compared to the previous literature on real-world feedback in tournaments.

First, we observe several dimensions of a manager's performance, and the managers are ranked on each of them. We exploit the fact that a manager's rank can be different on different dimensions to estimate the relationship between rank and change in performance. For example, if the manager does respond to his rank, we would expect profit to respond to profit rank, but not necessarily to service rank, or not in the same way; and vice versa. On the other hand, if the relationship between performance and rank is driven by something like working in spurts, then we might see it across the board: when effort has been low in the first 11 weeks, all ranks are high (bad) in week 11 and all measures increase faster in the last two weeks. So we may well see profit increase 'caused' by a bad rank on service. Hence, we can separate genuine response to rank on different dimensions of performance from more 'aggregate' behaviours such as working in spurts.

Second, whilst in many studies the key input is effort and is unobserved, we have data on some of manager's key input decisions. This allows to estimate the potential response to rank on variables that are considerably more directly under a manager's control than performance. This is useful for three reasons.

1. If we see effects of rank on both performance and directly controlled decisions, it is less likely that our performance results are due to patterns in random performance fluctuations, such as reversion to the mean.

2. Finding effects on managers' decisions makes it harder to explain away our results by general effort fluctuations, as in 'working in spurts' hypothesis.

3. In most cases, we know the theoretical relationship between managers' input decisions and performance (for example, higher output-labour ratio will increase profit but reduce service, as discussed in section 2). So, input data provide an important consistency test for performance results.

3.2 Data and patterns in the last two weeks

3.2.1 Data

We have data on key measures of manager performance (sales, profit and service) as well as total labour and output-labour ratio, which capture the key managerial decisions we discussed in section 2.1. Data span the period from 2010 to 2015. We exclude Q4 from the analysis, because due to winter holidays, the quarter end patters for Q4 are very different from other quarters. Hence, altogether we have 18 quarters of data. On average there are 230 shops per quarter, growing over time to 290. In total, the data include 530 individual managers.

The changes in our key variables in the last two weeks of the quarter have some big outliers, which we believe are driven by things like shop expansions or partial closures due to refurbishment. To deal with this, we drop the top and bottom 1% of outliers in all estimations.

3.2.2 End of quarter patterns

Ultimately, we are interested in whether there is heterogeneity in manager behaviour in the last two weeks before the tournament, depending on the feedback they got. But, as a first cut at the data, it is interesting to ask what happens in the last two weeks in aggregate. Table 3 shows that, on average, performance improves across all three measures, sales, profit and service (columns 1-3)[10]. This finding is consistent with other studies that have documented an improvement in performance towards the pay date (Oyer (1998), Kaur et al (2016)).

We also see that there is a significant change in the decisions under the manager's direct control: both labour and output-labour ratio rise (table 3, columns 4 and 5). This supports the idea that improvement in performance in the last two weeks is due to manager actions rather rather than external shocks. In section 4.1 we will see that these averages hide substantial heterogeneity. We can see some evidence of this heterogeneity already now: for each performance measure, table 4 shows that managers who improve performance by more improve their final tournament ranks by more[11]. The effects of input decision variables on ranks are these:

- Labour increases improve sales and service rank, but hurt profit rank
- Output-labour ratio increases improve sales and profit rank, but hurt service rank

These effects are consistent with the predictions we derived from our simple framework in section 2.1.

Although managers may improve their rank by improving performance this does not imply that they will increase their bonus. This is because bonus is paid in bands, and so only managers who are either near the boundary of the band or undertake a particularly large improvement in performance will increase their score and therefore bonus. Indeed, the data show that while *on average* managers who improve performance in the last two weeks do improve the score on that measure, these improvements are not very big (table 5[12]).

[10]The magnitudes of increases in terms of standard deviations are not large, but this is somewhat misleading because we have very different size shops in the sample, and so the standard deviation is very large due to cross-sectional differences.

[11]The regressions control for week 11 rank on each measure to deal with potential reversion to the mean.

[12]The regressions control for week 11 score on each measure to deal with potential reversion to the mean.

Table 3: Change in weeks 12-13 relative to weeks 1-11

	(1)	(2)	(3)	(4)	(5)
	Sales	Profit	Service	Labour	Output-labour ratio
Change	0.03	0.03	0.19	0.02	0.08
(in standard deviations)	(0.00)***	(0.00)***	(0.03)***	(0.00)***	(0.00)***
N	3,707	3,707	3,670	3,702	3,734

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 4: Change in rank in weeks 12-13

	Sales	Profit	Service
Increase in weeks 12-13 in:			
Underlying outcome	-0.01	-0.02	-0.09
	(0.00)***	(0.00)***	(0.00)***
R^2	0.23	0.30	0.58
N	3,218	3,187	3,213
Labour	-0.01	0.01	-0.02
	(0.00)***	(0.00)**	(0.00)***
R^2	0.22	0.28	0.24
N	3,220	3,190	3,190
Output-labour ratio	-0.01	-0.02	0.02
	(0.00)***	(0.00)***	(0.00)***
R^2	0.22	0.30	0.25
N	3,248	3,217	3,218

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager fixed effects and quarter effects, and corresponding week 11 rank

Table 5: Change in score in weeks 12-13

	Sales	Profit	Service
Increase in the last two weeks in:			
Underlying outcome	0.28	0.98	4.47
	(0.15)*	(0.16)***	(0.14)***
R^2	0.41	0.36	0.50
N	3,221	3,220	3,244
Labour	0.19	-0.26	0.76
	(0.15)	(0.16)*	(0.15)***
R^2	0.41	0.34	0.32
N	3,223	3,221	3,221
Output-labour ratio	0.18	1.11	-1.02
	(0.15)	(0.16)***	(0.16)***
R^2	0.41	0.35	0.32
N	3,251	3,249	3,249

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager fixed effects and quarter effects and corresponding score in week 11

For example, a one standard deviation increase in profit results in a 1 point improvement in the profit score on average. Since improvements in the score come in discrete steps, mostly 10 points high, this estimate is consistent with a scenario where by improving profit by one standard deviation the managers improve their pay one time in ten. Consistent with the idea that sales are subject to forces outside manager's control, return to improvements on sales is the lowest: 0.3 points for one standard deviation in improvement, only borderline significantly different from zero.

4 Results

4.1 Main results

4.1.1 Performance

We now look at the effects of the information contained in week 11 feedback on the change in the three performance measures - sales, profit and service - in weeks 12 and 13. Table 6 gives the results of estimating equation (8) for each measure.

Incentives

We first look at whether the managers who find themselves closer to a band boundary, and thus facing steeper marginal incentives in week 11 on a given measure, improve their performance by more. To do this, we focus on the top three rows of the table forming a 3 x 3 fragment (subsection of the table headed Incentives).The theory predicts positive and significant effects on the diagonal of this fragment.

In fact, what we see is that, with one exception, none of the incentive dummies are significant: the managers who find themselves near to the jump in bonus in week 11 do not behave differently from the rest. The only significant coefficient is a negative one on sales suggesting that managers who face steeper incentives on sales, reduce sales. This is the opposite of how we expect incentives to work.

Rank

Second, we examine the impact of ranks. Again, if we expect any significant coefficients, first of all they ought to be along the diagonal of the Rank section in table 6. Looking at the diagonal, while sales rank does not affect sales, profit rank and service rank both have significant effects on profit and service respectively. This shows that what managers are responding to is the feedback about their relative rank, while incentive effects are being controlled for.

Interestingly, these effects are positive: people with higher (worse) rank on a particular measure tend to improve their performance on that measure by more. Hence, we find evidence for 'desire to catch' up rather than 'discouragement'.

Recall that on average, managers improve performance in the last two quarters. Quantitatively, table 6 results tell us that when a manager gets the bottom rank instead of the top rank in week 11 (in a league table of 80, on average), he improves his profit in the last two weeks of the quarter by 0.3 standard deviation more. Similarly, if he finds he drops from top to bottom rank on service, he responds by improving service by 0.4 standard deviation more.

To sum up, we find that managers do not change their performance when they face steeper marginal incentives. However, they do respond to the rank they receive: They improve service and profit if they find themselves with a bad rank on these outcomes in week 11. Sales, which are harder to change in the short run, do not respond to rank.

4.1.2 Input decisions

The above interpretation of performance results is based on an assumption that the changes in performance we observe at least in part are due to some manager action. Indeed, we can examine manager decisions more directly by looking at labour and output-labour ratio, which are under the control of a manager to a considerably greater degree than performance measures. From our predictions in section 2.1, recall that typically increases in labour can improve sales and service, but will hurt profit. Furthermore, improvements in profit should be accompanied by a rise in output-labour ratio, while improvements in service – by a deterioration in the same ratio. This is useful interpreting results in table 7 which shows how labour and output-labour ratio respond to week 11 feedback.

Table 6: Outcomes

(Y = change in weeks 12-13)

	(1)	(2)	(3)
	Sales	Profit	Service
Incentives			
Sales	-0.09	-0.04	-0.04
	(0.04)**	(0.04)	(0.04)
Profit	-0.04	-0.01	-0.03
	(0.04)	(0.04)	(0.04)
Service	0.03	0.02	0.03
	(0.04)	(0.04)	(0.04)
Rank			
Sales	0.06	-0.04	0.08
	(0.07)	(0.08)	(0.07)
Profit	0.17	0.26	0.02
	(0.08)**	(0.09)***	(0.08)
Service	0.07	-0.00	0.39
	(0.07)	(0.07)	(0.14)***
Controls			
Incentives at the top	-0.02	0.01	0.06
	(0.07)	(0.07)	(0.07)
Level of Y	-0.06	-0.21	-0.48
	(0.04)	(0.04)***	(0.08)***
Shop size (proxy)	0.32	0.48	-0.02
	(0.08)***	(0.08)***	(0.06)
R^2	0.32	0.28	0.30
N	3,239	3,238	3,228

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Incentives

First, consider the effects of incentives (subsection Incentives of table 6). Higher marginal incentives on sales reduce labour (column 1) and improve output-labour ratio (column 2), which is the opposite of what we would expect, at least as far as labour is concerned. We do not have a good explanation for these effects. We find no other effects of either profit or service incentives on either labour or output-labour ratio (columns 1 and 2). So once again, we see that marginal incentives do not have expected effects, now on the variables the manager controls directly.

Rank

A different picture emerges when we look at the managers' response to rank (subsection Rank of table 7). First, consider the labour decision (column 1). When they rank badly on sales, managers increase labour, which makes sense if they are trying to increase sales. In contrast, managers who lag on profit cut labour, which is what section 2.1 predicts they would do if they want to improve profit. Thus this is consistent with evidence from table 6 that when lagging behind on profit, managers try to increase it. Finally, when managers rank low on service, we see that they increase labour. Again, this is consistent with our predictions about what the managers need to do to improve service and hence with the evidence in table 6 that managers who rank poorly on service respond by improving it.

Second, consider output-labour ratio (column 2). Recall that the sign of the relationship between sales and output-labour ratio is ambiguous, and also sales may be hard to shift in the short run. So perhaps it is not surprising that we do not find any impact of sales rank on output-labour ratio. At the same time, managers who have received a bad rank on profit in week 11 improve output-labour ratio by more, which is in line with our prediction about what should happen if they try to improve profit. In contrast, managers who face a bad rank on service, do the opposite: they allow output-labour ratio to deteriorate; this is again consistent with our predictions and the idea that

Table 7: Input decisions

(Y = change in weeks 12-13)

	(1)	(2)
	Labour	Output-labour ratio
Incentives		
Sales	-0.16	0.11
	(0.04)***	(0.04)***
Profit	-0.05	0.04
	(0.04)	(0.04)
Service	-0.00	0.04
	(0.04)	(0.04)
Rank		
Sales	0.20	-0.11
	(0.08)**	(0.07)
Profit	-0.27	0.39
	(0.09)***	(0.08)***
Service	0.34	-0.36
	(0.07)***	(0.07)***
Controls		
Incentives at the top	-0.14	0.01
	(0.07)*	(0.07)
Level of Y	-0.04	-0.28
	(0.03)	(0.04)***
Shop size (proxy)	0.12	0.07
	(0.08)	(0.06)
R^2	0.28	0.21
N	3,239	3,269

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects these managers are trying to improve service.

Hence our evidence on managers' input decvisions supports our earlier evidence from performance measures. First, managers generally do not respond to marginal incentives. Second, they respond to rank. The results from profit and service are clear cut: managers who get worse rank on these measures try to improve them and adjust their input decisions in ways that help deliver this. The results on sales are weaker, and less clear cut. This perhaps is unsurprising given that sales are harder for managers to influence, at least in the short run.

4.2 Career concerns

Our results above show that managers do not respond to marginal incentives of the tournament, but instead they react to rank and try to improve performance on the measures where their rank is worse.

If the firm's retention and promotion policies are related to the rank that the managers receive in quarterly tournaments, our results can be due to managers responding to such career concerns. In this section we show that this is not the case.

We start by asking whether career prospects at this firm are indeed linked to tournament ranks (table 8). We look at whether there is a relationship between these ranks in a manager's early career and the probability of staying longer with the firm (columns 1 and 2) as well as the probability of being promoted to the next level (columns 3 and 4)[13,14].

In both cases, we find a correlation with profit rank which runs in the expected direction, but no correlation with other two ranks (columns 1 and 3). The correlation with profit rank disappears when we control for the overall

[13] We estimate a linear probability model. In columns 1 and 2 we look at whether ranks in the first 1.5 years are correlated with the probability of the manager staying on with the firm for 2.5 years or more; our results are the same if we use other cut offs. In our data, the average retention rate beyond 2.5 years for managers who have spent at least 1.5 years with the firm is 78%. This is an underestimate, because our data cannot distinguish between departure from the firm and a manager being promoted beyond the next level of the hierarchy.

[14] On average, 7% of managers get promoted to the next level.

Table 8: Impact of rank on career prospects

	(1)	(2)	(3)	(4)
	Probability of			
	Departure	Departure	Promotion	Promotion
Sales rank	-0.06	-0.19	-0.06	0.09
	(0.08)	(0.11)*	(0.05)	(0.07)
Profit rank	0.14	0.02	-0.16	0.00
	(0.08)*	(0.11)	(0.05)***	(0.07)
Service rank	0.02	-0.08	-0.03	0.10
	(0.10)	(0.12)	(0.06)	(0.08)
Overall rank		0.29		-0.34
		(0.16)*		(0.11)***
Total experience			0.01	0.01
			(0.00)***	(0.00)***
Constant	0.17	0.20	0.14	0.10
	(0.07)**	(0.08)***	(0.05)***	(0.05)*
R^2	0.01	0.01	0.06	0.07
N	506	496	600	597

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

tournament rank, which has a significant negative relationship with retention and promotion (lower, i.e. better, ranks are more likely to stick around and to be promoted). Hence, although manageres who have a better rank do indeed face better prospects with the firm, unsurprisingly, the relationship is between overall rank rather than ranks on individual measures.

It is therefore possible that our main results on individual performance measures (tables 6 and 7) are in fact driven by the managers' desire to improve overall rank in response to career incentives at the firm. To check this, we estimate our original regressions now controlling for the overall rank that the manager receives in week 11 (tables 11 and 12 in Appendix). If this explanation is correct, the ranks on individual measures should no longer be significant. Instead, we find that the results on both performance outcomes (table 11) and inputs (table 12) are unchanged.

4.3 Non-linear effects

We have established that managers who receive worse week 11 rank on profit and service try to catch up by improving them, and make decisions about labour and output-labour ratio to deliver those improvements. So far, we used a linear model to estimate these effects. But it is possible that the effects are non-linear and even non-monotonic (as in Gill et al (2016)). Do managers at the top display particular complacency, or instead might they work harder encouraged by their top position? Conversely, do managers at the bottom get discouraged and drop out or do they work harder still to avoid being last?

To explore these questions, we re-estimate our main equation (8) adding dummy variables for whether a manager was in top 5% or bottom 5% of the rank on each of the tournament measures, sales, profit and service.

In table 9 we look at the impact of rank on change in outcomes. Fundamentally, our results are unchanged: we continue to see the positive effects of the linear rank variable on profit and service outcomes we have found before. However, alongside these we now see additional complacency at the top in profit measure and additional dislike of being at the bottom in the service measure.

When we look at the effects of rank on manager controlled variables, labour and output-labour ratio (table 10), we again see that linear rank effects remain the same in sign and significance and similar in magnitude to those in the main regressions of table 6. At the same time, there is evidence of additional dislike of being at the bottom in profit/labour relationship, but no non-linear effects in service.

In summary, the patterns that emerges, although not exactly the same across different measures, allow us to say three things. First, the effect of rank is monotonic - in all specifications where rank effects are present, desire to improve always increases with rank. Second, it is not driven purely by the bottom ranks. Third, what evidence of non-linearity there is points at particular complacency at the top and dislike of bottom ranks.

4.4 Robustness checks

In this section we undertake two robustness checks. First, we repeat the estimations only for bands that are relatively wide. Second, we use alternative definitions of how close the manager needs to be to the boundary to face steep incentives.

4.4.1 Wider bands

The number of distinct ranks in a band varies between 2 and 47, with the average of 14. The variation comes from three factors: first, more ties in performance means less distinct ranks in a band; second, the firm deliberately makes bands in the middle wider than bands on the edge; third, the number of regions into which shops are divided has varied over time between 2 and 4, and when there are more regions, the bands have to be narrower.

Since we define our incentive dummy as being within two ranks of the next band (up or down), narrower bands will contribute a relatively bigger share of observations to treatment group than to control group compared to wider

Table 9: Non-linear effects - outcomes

(Y = change in weeks 12-13)

	(1)	(2)	(3)
	Sales	Profit	Service
Sales			
Rank	0.10	0.02	0.03
	(0.08)	(0.09)	(0.08)
Top 5%	0.13	0.13	-0.02
	(0.11)	(0.11)	(0.10)
Bottom 5%	-0.03	-0.08	0.16
	(0.09)	(0.10)	(0.09)*
Profit			
Rank	0.13	0.20	0.01
	(0.09)	(0.09)**	(0.09)
Top 5%	-0.28	-0.31	-0.17
	(0.10)***	(0.11)***	(0.10)*
Bottom 5%	-0.02	0.02	-0.08
	(0.09)	(0.10)	(0.09)
Service			
Rank	0.07	0.01	0.44
	(0.08)	(0.08)	(0.14)***
Top 5%	0.06	0.06	-0.08
	(0.09)	(0.09)	(0.09)
Bottom 5%	0.06	0.01	0.19
	(0.09)	(0.09)	(0.09)**
R^2	0.32	0.28	0.30
N	3,239	3,238	3,228

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives, shop size, level of Y in week 11

Table 10: Non-linear effects - input decisions

(Y = change in weeks 12-13)

	(1)	(2)
	Labour	Output-labour ratio
Sales		
Rank	0.20	-0.06
	(0.09)**	(0.08)
Top 5%	0.16	-0.04
	(0.11)	(0.10)
Bottom 5%	0.13	-0.20
	(0.10)	(0.09)**
Profit		
Rank	-0.22	0.36
	(0.09)**	(0.09)***
Top 5%	-0.04	-0.12
	(0.11)	(0.10)
Bottom 5%	-0.20	0.06
	(0.10)**	(0.09)
Service		
Rank	0.32	-0.35
	(0.08)***	(0.08)***
Top 5%	-0.02	-0.02
	(0.09)	(0.09)
Bottom 5%	0.06	-0.07
	(0.09)	(0.09)
R^2	0.28	0.22
N	3,239	3,269

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives, shop size, level of Y in week 11 bands. This can lead to biased estimates if the observations in the narrow bands have different underlying values of the dependent variable, in absence of any incentive effects. The direction of bias will depend on the sign of this correlation.

Although the mechanisms that would lead to such correlation are not obvious, in order to reduce the potential effects of this, we re-run all the regressions defining incentive dummies in the same way as before, but only for bands that are at least 7 ranks wide. This means that about 1/3 of observations is dropped. Table 13 in Appendix shows that we continue to get no effect of marginal incentives on outcomes. At the same time, the impact of rank on the outcomes remains the same. Table 14 in Appendix further shows that our original results on labour and output-labour ratio are also unaffected.

4.4.2 Alternative definition of incentives

While the theory of tournaments tells us that managers close to the boundary of the bonus band will face steeper incentives, it does not give us much of a guide for how close is close enough. So far, we have defined a manager as facing steeper incentives if he is in a two rank window on either side of the boundary. To check whether our results are robust to alternative definitions, we re-run the estimations with windows of 1 rank (not reported) and 3 ranks (tables 15 and 16 in Appendix). Our results are unchanged.

5 Conclusions

We have analyzed how store managers who face a high stakes tournament incentive scheme react to regular feedback on three key performance measures. We see three main take aways from our results:

1. Marginal incentives, even when sizeable, are rather ineffective in improving performance

2. Providing people with rank is effective: on average, people who get worse rank try to catch up. This suggests that on average, desire to catch up

dominates potential discouragement effects from getting a bad rank.

3. In a multi-tasking setting, providing rank for each measure can serve to achieve improvements on performance aspects that lag behind.

Interestingly, in a setting like ours, where incentives and rank impacts act on different managers (forming the basis for our empirical identification), rank effects lead to a different distribution of performance improvements that could be expected from incentive effects. These differences in distribution could potentially be important to the firm, and deliver different aggregate results.

Crucially, our analysis is done *on the margin*, so our results do not suggest that a firm will do better by scrapping a tournament incentive scheme and merely providing a ranking. It is possible that monetary and non-monetary motivations interact in complex ways; this is an important and somewhat under-explored direction for research.

Another important point concerns our finding that when faced with a bad rank, on average people choose to catch up rather than to stop trying. In our study this is conditional on the participation decision of the manager. On the other hand, discouragement may well operate through the participation decision: a question that we are planning to explore in future work.

References

Aoyagi, Masaki (2010) "Information feedback in a dynamic tournament", *Games and Economic Behaviour*, Vol. 70, pp. 242-260.

Azmat, Ghazala and Nagore Iriberri (2015) "The Provision of Relative Performance Feedback: An Analysis of Performance and Satisfaction", *Journal of Economics and Management Strategy*, Vol. 25, No. 1, pp. 77-110.

Azmat, Ghazala, Manuel Bagues, Antonio Cabrales and Nagore Iriberri (2016) "What You Dont Know... Cant Hurt You? A Field Experiment on Relative

Performance Feedback in Higher Education", IZA Discussion Papaper No. 9853.

Casas-Arce, Pablo and Asís Martínez-Jerez (2009) "Relative Performance Compensation, Contests, and Dynamic Incentives", *Management Science*, Vol. 55, No. 8, pp. 1306-1320.

Barankay, Iwan (2012) "Rank Incentives: Evidence from a Randomized Workplace Experiment", mimeo, University of Pennsylvania.

Blanes i Vidal, Jordi and Mareike Nossol (2011) "Tournaments without prizes: evidence from personnel records", *Management Science*, Vol. 57, No. 10, pp. 1721-1736.

Delfgaauw, Josse, Robert Dur, Arjan Non, and Willen Verbeke (2014) "Dynamic incentive effects of relative performance pay: A field experiment", *Labour Economics*, Vol. 28, pp. 1-13.

Ederer, Florian (2010) "Feedback and Motivation in Dynamic Tournaments", *Journal of Economics and Management Strategy*, Vo. 18, No. 3, pp. 733-769.

Ehrenberg, Ronald G. and Michael L. Bognanno (1990) "Do Tournaments Have Incentive Effects?", *Journal of Political Economy*, Vol. 98, No. 6, pp. 1307-1324.

Eriksson, Tor, Anders Poulsen, and Marie Claire Villeval (2009) "Feedback and incentives: Experimental evidence", Vol. 16, pp. 679-688.

Genakos, Christos and Mario Pagliero, "Interim Rank, Risk Taking, and Performance in Dynamic Tournaments", *Journal of Political Economy*, Vol. 120, No. 4, pp. 782-813.

Gill, David, Zdenka Kissová, Jaesun Lee, and Victoria Prowse (2015) "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision", IZA Discussion Papaper No. 9286.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan (2015) "Self-Control at Work", *Journal of Political Economy*, Vol, 123, No. 6, pp. 1227-1277.

Lazear, Edward P., and Sherwin Rosen (1981) "Rank-order tournaments as optimum labour contracts", *Journal of Political Economy*, Vol. 89, No. 5, pp. 841-864.

Müller, Wieland and Andrew Schotter (2009) "Workaholics and dropouts in organizations", *Journal of the European Economic Association*, Vol. 8, No. 4, pp. 717-743.

Nalebuff, Barry J, and Joseph E. Stiglitz (1983) "Prizes and incentives: Towards a general theory of compensation and competition", *The Bell Journal of Economics*, Vol. 14, No. 1, pp. 21-43.

Oyer, Paul (1998) "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality", *The Quarterly Journal of Economics*, Vol. 13, No. 1, pp. 149-185

Appendix: Robustness checks

Table 11: Outcomes, controlling for overall rank
(Y = change in weeks 12-13)

	(1)	(2)	(3)
	Sales	Profit	Service
Incentives			
Sales	-0.09	-0.04	-0.04
	(0.04)**	(0.04)	(0.04)
Profit	-0.04	-0.01	-0.03
	(0.04)	(0.04)	(0.04)
Service	0.03	0.02	0.03
	(0.04)	(0.04)	(0.04)
Rank			
Sales	0.07	-0.02	0.13
	(0.10)	(0.10)	(0.10)
Profit	0.19	0.28	0.08
	(0.12)	(0.12)**	(0.11)
Service	0.10	0.02	0.45
	(0.12)	(0.12)	(0.16)***
Overall rank	-0.04	-0.04	-0.12
	(0.17)	(0.18)	(0.17)
R^2	0.32	0.28	0.30
N	3,239	3,238	3,228

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size

Table 12: Labour decisions, controlling for overall rank

(Y = change in weeks 12-13)

	(1)	(2)
	Labour	Output-labour ratio
Incentives		
Sales	-0.16	0.11
	(0.04)***	(0.04)***
Profit	-0.05	0.04
	(0.04)	(0.04)
Service	-0.00	0.04
	(0.04)	(0.04)
Rank		
Sales	0.15	-0.02
	(0.10)	(0.10)
Profit	-0.33	0.50
	(0.12)***	(0.11)***
Service	0.28	-0.23
	(0.12)**	(0.11)**
Overall rank	0.11	-0.24
	(0.18)	(0.17)
R^2	0.28	0.21
N	3,239	3,269

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size

Table 13: Wider bands - outcomes

(Y = change in weeks 12-13)

	(1)	(2)	(3)
	Sales	Profit	Service
Incentives			
Sales	0.02	0.05	-0.01
	(0.05)	(0.05)	(0.05)
Profit	0.00	0.01	-0.03
	(0.05)	(0.05)	(0.05)
Service	0.05	0.03	0.07
	(0.04)	(0.04)	(0.04)
Rank			
Sales	0.08	0.03	0.06
	(0.09)	(0.09)	(0.09)
Profit	0.21	0.35	-0.09
	(0.10)**	(0.11)***	(0.10)
Service	0.12	0.02	0.49
	(0.08)	(0.08)	(0.17)***
R^2	0.44	0.39	0.33
N	2,254	2,249	2,234

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size

Table 14: Wider bands - input decisions

(Y = change in weeks 12-13)

	(1)	(2)
	Labour	Output-labour ratio
Incentives		
Sales	-0.10	0.17
	(0.05)**	(0.05)***
Profit	0.01	0.04
	(0.05)	(0.05)
Service	0.01	0.02
	(0.05)	(0.04)
Rank		
Sales	0.20	-0.09
	(0.10)**	(0.09)
Profit	-0.36	0.47
	(0.11)***	(0.10)***
Service	0.49	-0.42
	(0.09)***	(0.08)***
R^2	0.36	0.27
N	2,252	2,268

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size

Table 15: Incentives within 3 ranks - outcomes

(Y = change in weeks 12-13)

	(1)	(2)	(3)
	Sales	Profit	Service
Incentives			
Sales	-0.05	-0.03	-0.01
	(0.04)	(0.04)	(0.04)
Profit	-0.02	0.00	0.00
	(0.04)	(0.04)	(0.04)
Service	0.09	0.07	0.01
	(0.04)**	(0.05)	(0.04)
Rank			
Sales	0.05	-0.04	0.08
	(0.07)	(0.08)	(0.07)
Profit	0.17	0.26	0.02
	(0.08)**	(0.09)***	(0.08)
Service	0.06	-0.01	0.40
	(0.07)	(0.07)	(0.14)***
R^2	0.32	0.28	0.30
N	3,239	3,238	3,228

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size

Table 16: Incentives within 3 ranks - input decisions

(Y = change in weeks 12-13)

	(1)	(2)
	Labour	Output-labour ratio
Incentives		
Sales	-0.11	0.10
	(0.04)***	(0.04)***
Profit	-0.06	0.06
	(0.04)*	(0.04)
Service	0.03	0.05
	(0.05)	(0.04)
Rank		
Sales	0.19	-0.12
	(0.08)**	(0.07)
Profit	-0.27	0.38
	(0.09)***	(0.08)***
Service	0.34	-0.37
	(0.07)***	(0.07)***
R^2	0.28	0.21
N	3,239	3,269

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Dependent variable is in standard deviations

With manager and regional fixed effects and quarter effects

Controls: incentives at the top, level of Y, proxy for shop size