



TI 2009-069/1

Tinbergen Institute Discussion Paper

Tournament Incentives in the Field: Gender Differences in the Workplace

Josse Delfgaauw^{1,2}

Robert Dur^{1,2,3,4}

Joeri Sol^{1,2}

Willem Verbeke^{1,5}

¹ Erasmus University Rotterdam;

² Tinbergen Institute;

³ CESifo;

⁴ IZA;

⁵ ERIM.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

Tournament Incentives in the Field: Gender Differences in the Workplace*

Josse Delfgaauw[†], Robert Dur[‡], Joeri Sol[§], and Willem Verbeke[¶]

31 July 2009

Abstract

We ran a field experiment in a Dutch retail chain consisting of 128 stores. In a random sample of these stores, we introduced short-term sales competitions among subsets of stores. We find that sales competitions have a large effect on sales growth, but only in stores where the store's manager and a large fraction of the employees have the same gender. Remarkably, results are alike for sales competitions with and without monetary rewards, suggesting a high symbolic value of winning a tournament. Lastly, despite the substantial variation in team size, we find no evidence for free-riding.

JEL-codes: C93, J16, M52.

Keywords: field experiment, gender differences, competition, sales contests, awards.

*We gratefully acknowledge comments by Martin Bøg, Alison Booth, Imran Rasul, and seminar audiences at Universitat Autònoma de Barcelona, Bar-Ilan University, Erasmus University Rotterdam, and the University of Zürich, and participants to the 20th Jerusalem School in Economic Theory at the Hebrew University.

[†]Erasmus University Rotterdam and Tinbergen Institute. E-mail: delfgaauw@ese.eur.nl.

[‡]Erasmus University Rotterdam, Tinbergen Institute, CESifo, and IZA. E-mail: dur@ese.eur.nl.

[§]Erasmus University Rotterdam and Tinbergen Institute. E-mail: jsol@ese.eur.nl.

[¶]Erasmus University Rotterdam and ERIM. E-mail: verbeke@ese.eur.nl.

1 Introduction

Throughout the world, in business as well as in government, men are strongly overrepresented in top positions. For instance, in 2008, only 16% of all ministerial positions worldwide were held by women; similarly, among the world's 192 heads of government, there were only eight women (IPU 2008). In business, the situation is not much different. For example, in a large sample of publicly traded US firms, Bertrand and Hallock (2001) find that only 2.5% of the five highest-paid positions are held by women. Wirth (2001) reports similar patterns for other countries.

Traditional explanations for the small number of women in top positions are occupational sorting resulting from gender differences in ability or preferences (Polachek 1981) and gender discrimination (e.g. Snizek and Neil 1992). Inspired by evolutionary biology, recent experimental studies – starting with Gneezy et al. (2003) – suggest a third explanation: men are more strongly motivated by competitive incentives or more effective in competitive environments than women, thus impeding women in competitions for promotions or for new jobs.

By now, there is quite some empirical support for such gender differences. In a lab experiment, Gneezy et al. (2003) let participants solve computerized mazes and varied the competitiveness of the environment. They find that, while men and women perform equally well under individual piece rates, men perform much better than women under competitive incentives. Gneezy and Rustichini (2004) show that these gender differences are already present at a very young age. In a 40 meter dash, nine-year-old boys run much faster in a race than when they run alone. By contrast, while girls run as fast as boys when running alone, competition does not increase their running speed. In non-experimental settings, underperformance of women under competitive pressure is found in student admissions to schools (Jurajda and München 2008 and Örs et al. 2008) and in Grand Slam tennis (Paserman 2007). The recent field study by Lavy (2008), however, finds no gender differences in the effect of relative performance pay on high-school teacher's performance in Israel. Croson and

Gneezy (2008) provide a recent overview of the literature.¹

This paper studies gender differences in competition by conducting a field experiment in a naturally occurring work environment. A unique feature of our analysis is that we study competition among teams of employees, each headed by a professional manager. Using the variation in the gender composition of the teams as well as in the gender of the manager, we test whether female-dominated and female-led teams respond differently to competitive incentives, which were introduced in a random sample of the geographically dispersed teams. Moreover, we test for possible interaction effects between the gender of the manager and the gender composition of the team. Studying gender differences in competition among manager-led teams is most relevant in the context of the sharp gender differences in holding executive-level positions discussed above. Reaching an executive-level position, be it in business or government, commonly requires winning several promotion or job competitions. These competitions are often decided by candidates' relative performance which (except for employees at the lowest hierarchical level) depends not merely on one's own effort or talent, but also crucially on the performance of the members of the team one leads.

More concretely, we ran a field experiment in a discount retail chain in The Netherlands specializing in shoes, sports apparel, and casual clothing. In a randomly selected subset of the 128 stores, we introduced short-term sales competitions among stores. The selected stores were divided into pools of 5 and competed for a period of 6 weeks on the basis of percentage sales growth compared to the same period the year before. Each week, the stores that took part in a competition received a poster with cumulative sales growth figures of the stores in their pool, ranked in descending order. After the sales competitions, the store manager and all employees of the store with the highest sales growth in a pool received a bonus of 75 euro; all employees and the store manager of each pool's runner-up received 35

¹A closely related strand in the experimental literature studies self-selection into competitive environments. Datta Gupta et al. (2005), Dohmen and Falk (2006), and Niederle and Vesterlund (2007) find that men opt significantly more often for competitive compensation schemes than women. Gneezy et al. (2008) show that the reverse holds in a matrilineal society. Recent studies have shown that the gender gap in self-selection into competition by and large vanishes for girls attending single-sex schools (Booth and Nolen 2009) and when the tournament is among teams rather than among individuals (Dargnies 2009).

euro.

We find that, on average, the tournaments increase weekly sales growth by about five percentage points. More interestingly, we find no significant difference in the effect of tournaments on sales growth between stores with a male manager and stores with a female manager, nor do we find that sales competitions have a larger effect on performance in stores with a higher fraction of male employees. However, we do find a remarkable interaction effect of these two gender variables on sales growth responsiveness: in stores with a male manager, the effect of competition increases in the share of male employees, while the reverse holds for female-led stores. These effects are substantial: when a male manager’s team has a sufficiently high fraction of male employees, there is a large and highly significant effect of sales competitions, while there is no significant effect when the fraction of female employees is high. By contrast, in stores with a female manager, the point estimate of the effect of competition on performance is insignificantly negative when the share of male employees is high, but turns positive when a sufficiently high fraction of employees is female.

A number of recent studies argue that competition can motivate people not merely because of the chance of winning a monetary reward, but also because of non-pecuniary benefits such as perceived esteem, status, and social recognition (Auriol and Renault 2008, Besley and Ghatak 2008, Frey and Neckermann 2008, Moldovanu et al. 2007). Neckermann and Kosfeld (2008) show experimentally that a tournament with no more at stake than an award of zero material value can have a great impact on people’s performance. Likewise, Blanes i Vidal and Nossol (2009) and Azmat and Iriberry (2009) find that simply providing information to subjects about their relative performance boosts performance substantially. Bandiera et al. (2009), however, find the opposite effect. In our experiment, parallel to the treatment described above, another subset of stores competed in tournaments with the same setup except for the absence of a monetary reward for winning. So, stores also competed in pools of five, for a period of six weeks, and received a weekly ranking of stores in their pool based on sales growth, but neither the manager nor the employees could earn a bonus. Remarkably,

we find that tournaments without monetary rewards have a significantly positive effect on sales growth. The effect is of similar magnitude to the effect of tournaments with monetary rewards, suggesting a high symbolic value of winning a tournament. Gender differences in the effects of competition are also similar in both treatments.

The key finding of our study is that competition among manager-led teams can be a powerful motivator for men as well as for women, but only when the manager and a large fraction of the employees have the same gender. We can think of two plausible interpretations of this result. First, the response of team members to competition may crucially depend on the way a competition is communicated and promoted by the team’s manager. In our experiments, we deliberately left a lot of discretion to team managers on how to use the competition as an incentive device. In particular, both the announcement of the competitions and the weekly posters were only sent to the store managers, not to the employees. It was up to the team managers to make the competitions appealing to their employees. Our results suggest that both male and female managers succeeded in making the competition appeal to employees of their own sex, but failed to do so to employees of the opposite sex.² This is well in line with evidence from management studies showing that when working for a manager of the opposite sex, employees find their duties and responsibilities much more ambiguous than when working for a manager of the same sex (Tsui and O’Reilly 1989 and McNeilly and Russ 2000). Relatedly, a number of studies in organizational psychology have shown that, as compared to employees with opposite-sex managers, those with same-sex managers are more likely to develop high-quality leader-member exchange relationships (LMX) – a widely used measure of manager-employee mutual support, trust, and obligation – which may in turn facilitate communication.³ Lastly, experimental evidence using a subject pool of both students and banking executives finds that female participants tend to feel more comfortable supervising a female person than a male person in a challenging task, while male participants

²The idea that dissimilarity in personal attributes such as gender can deteriorate communication in organizations dates back to at least March and Simon (1958), who argue that dissimilarity may give rise to ‘language incompatibility’ and less frequent communication.

³See e.g. Duchon et al. (1986), Pelled and Xin (2000), and Varma and Stroh (2001). Wayne et al. (1994) discuss a number of reasons for why these differences may arise.

expect fewer conflicts with a male subordinate and perceive males to be more competent in a challenging task (Mai-Dalton and Sullivan 1981).

Alternatively, since our experiments involve competition among teams, the team nature of the incentive scheme may drive part of the differences in responsiveness between differently composed teams. Thus, managers may have no difficulties in making the competition appeal to employees of the opposite sex, but a male (female) manager may be better in strengthening the team’s internal cohesion if many team members are male (female). As a result, free-rider problems should be less severe in stores where the manager and a large part of the store’s employees have the same gender.⁴ Using the variation in the size of stores, we test for free-riding behavior of employees. However, despite the substantial variation in team size in our sample, we find no evidence that bigger stores respond less to sales competitions, neither in the treatment with bonus nor in the treatment without bonus. Moreover, we find no evidence that free-riding is less severe when the manager and a large part of the employees have the same gender.

We proceed as follows. In the next section, we describe the experimental set-up and the data. Section 3 describes the methodology of our empirical analysis and Section 4 reports the results. Section 5 concludes.

2 Experimental set-up and data description

The field experiment took place in 2007-2008 in a discount retail chain in The Netherlands, selling male and female clothing, shoes, and sports apparel. The chain consists of 128 geographically dispersed stores operating under one brand name and employing a total of 1574 people. Store employees earn an hourly wage slightly above the legal minimum hourly wage. Store managers earn about 45% more and their pay partially depends on their store’s performance.⁵ The company’s management wished to intensify the use of incentives. In

⁴Free-rider problems may arise because in stores employing relatively many employees, each individual employee has a smaller effect on the store’s percentage sales growth compared to colleagues in smaller stores.

⁵On average, slightly less than 5% of manager’s earnings is performance-related.

consultation with the management, we designed sales competitions among subsets of stores, with stores' percentage sales growth compared to the same period last year as the performance measure. Percentage sales growth is a commonly used performance measure in this company and is one of the key determinants of store managers' performance pay. We decided to introduce relative performance incentives rather than incentives based on absolute targets, as sales are very volatile (see Figure 1). A large part of this volatility is caused by common shocks (weather, holidays, advertising campaigns on national television, etc.),⁶ which renders relative performance pay attractive (Lazear and Rosen 1981, Green and Stokey 1983, and Nalebuff and Stiglitz 1983).

In the sales competitions, stores competed in pools of five during a period of six weeks. Stores received weekly feedback in the form of a poster containing cumulative sales growth figures for all five stores in their pool, ranked in descending order. Store managers were instructed to put up these posters in the store's canteen, where employees drink coffee and have lunch. The posters as well as the instructions were sent to the store managers through the company's usual channels; store managers and store employees did not know they took part in an experiment. Hence, our experiment can be classified as a natural field experiment (Harrison and List 2004). Store employees were not informed about the sales competitions by the company's management; it was up to the store managers to promote the competition.

Our study comprises two experimental treatments and an untreated control. First, in the 'bonus' treatment, stores compete for a monetary reward. The store manager and all employees of the winning store and of the runner-up received a reward of 75 and 35 euros, respectively.⁷ Second, in the 'feedback' treatment, no monetary reward could be won. Apart from the presence or absence of a monetary reward, the bonus treatment and the feedback treatment were identical.

Our dataset covers a period of 84 weeks (starting in week 1 of 2007). Sales competitions

⁶A panel regression including only week fixed effects explains about 65% of the variation in stores' sales growth.

⁷The first prize was about 5 percent of an employee's monthly wage. Rewards were halved for part-time employees.

took place in two periods of 6 weeks (in weeks 44 - 49 and weeks 71 - 76). Figure 2 gives an overview of all the events related to the experiment. In the first experimental period, all stores were assigned to one of the two experimental treatments, either bonus or feedback. In the second period, we included a control group of stores not taking part in a competition. We decided against a control group in the first period, because at that time we intended to focus our study on the effects of monetary rewards in tournaments. The second experimental period gives us the opportunity to also assess the effects of tournaments per se.

Competition provides stronger incentives when contestants are more homogeneous (Lazear and Rosen 1981). Bearing this in mind, we used data on past sales performance to create relatively homogeneous pools of stores. The assignment procedure for the first wave of the experiment was as follows. All stores were ranked according to percentage sales growth over the weeks 1 up to 37 compared to the same period the year before. The five stores with the highest sales growth were grouped into one pool and assigned to the bonus treatment; the next five were grouped into the next pool and assigned to the feedback treatment. This process was iterated consecutively until all 125 stores were grouped into a pool and allocated to either one of the treatments.⁸

In the second experimental period the assignment procedure was partly imposed by the company. For fairness reasons, the company obliged us to assign all stores who were in the feedback treatment during the first period to the bonus treatment in the second period.⁹ We grouped these stores into new pools of five stores each, this time using sales performance in weeks 50 to 68 to create relatively homogeneous pools. The remaining stores were assigned either to the feedback condition or to the untreated control group according to a similar procedure as before: the five best-performing stores in weeks 50 to 68 were assigned to the control condition, the next five stores were grouped into a pool and became part of the feedback condition, and so on. To avoid confusion and diminish sabotage opportunities, we

⁸During the first round of the experiment 3 stores were closed for renovation.

⁹The company wished, at a later point in time, to evaluate the experiment together with the store managers and feared that it would be considered unfair when some stores had never been assigned to the bonus condition.

replaced a store when two stores from the same city happened to be assigned to the same pool. In both periods, we made two of these adjustments.

The company provided us with the weekly sales data of each store, presented in indexes for confidentiality reasons. We used these to calculate weekly sales growth compared to the same week a year earlier. We also received each store’s personnel file before both experimental rounds, with information on gender, age, and tenure of the store’s manager and employees. Descriptive statics are given in Table 1. Across all stores, average weekly sales growth was negative in the period we consider. The retail chain had slightly less female-led stores than male-led stores.¹⁰ The average store had 12 employees (excluding the store manager), of which 85% was female. The average age and tenure of store managers was 39 years and 12 years, respectively. Some of the stores underwent a renovation, which made their appearance more modern without changing the range of products sold. Before the first experiment, six stores had been renovated; at the start of the second experimental period there were 17 renovated stores. As stores are closed during renovation, there are 268 missing store-week observations. In the analysis we control for the effects of renovation on subsequent sales growth by including a dummy variable which takes value 1 from the week in which a renovated store is reopened onwards.

Table 1 also reports the descriptive statistics within the two treatment groups and the control group to which stores were assigned in the second experimental period (see Figure 2). The three groups of stores hardly differ on observables. A randomization check using F-tests reveals that there are no statistically significant differences in the means of the observables between the three groups. Table 2 reports the descriptive statistics separated by store managers’ gender, where we only include the 114 stores we use in analyzing the gender effects (see footnote 10). Over the whole period, male-managed stores reached 0.4

¹⁰In both personnel files, for some stores information about the manager is missing, either because the store temporarily had no manager, or (in a single case) a store had two managers. In one store, a male manager was replaced by a female manager in between the two experimental periods. When analysing gender effects, we exclude these stores from the analysis. This leaves 114 stores. In five other stores, the manager was replaced by a manager with the same gender; excluding these stores from the analysis does not affect the results qualitatively.

percentage points higher weekly sales growth than female-managed stores, but the difference is not statistically significant. Male managers have significantly longer tenure than female managers, and run stores with significantly more employees. In the analysis below, we perform robustness checks where we control for these differences. Importantly, there is quite a lot of variation in the percentage of female employees, both for male- and female-led stores. Figure 3 depicts the distribution of the percentage of female employees in stores, separated by managers’ gender. Gender of the manager is not significantly related to the gender composition of store employees. Note that there are no stores with a majority of male employees. This implies that our estimates of the effect of stores’ gender composition are based on, and, hence, relevant for female-dominated teams.

3 Method

We estimate the effects of the competitions on sales growth using OLS panel estimation including week and store fixed effects. Let y_{st} be the sales growth of store s in week t . Further, let B_{st}^1 be a dummy variable which is equal to one during the *first* experimental period when store s was assigned to the bonus treatment (rather than to the feedback treatment) in the first experimental period. Similarly, let B_{st}^2 and F_{st}^2 be dummy variables for whether in the *second* experimental period, store s was assigned to the bonus treatment and to the feedback treatment respectively (rather than to the control group). To assess the average effect of the treatments in both experimental periods, we estimate:

$$y_{st} = \alpha_s + \theta_t + \gamma B_{st}^1 + \delta B_{st}^2 + \mu F_{st}^2 + \nu R_{st} + \varepsilon_{st} \tag{1}$$

where α_s and θ_t are store fixed effects and week fixed effects, respectively, R_{st} is a dummy for whether store s had been renovated before week t , and ε_{st} is an error term.

Observe that we allow the effect of the bonus treatment relative to the feedback treatment to differ between the first and second experimental period, i.e. we do not restrict that $\gamma =$

$\delta - \mu$. Loosely speaking, for each experimental period, we estimate difference-in-difference effects of the treatments, where we assume that in all non-experimental weeks, it is ‘business-as-usual’ for all stores. Hence, we do not allow for carry-over effects of treatments into the weeks following an experimental period. We have checked the robustness of this approach in two ways. First, none of our results is affected qualitatively if we exclude the first 8 weeks after either experimental period (weeks 50 - 57 and/or weeks 77 - 84) from our analysis. Second, all our results from the first experimental period carry over to an estimation which includes only the first 49 weeks (i.e. which excludes all weeks after the first period; see Figure 2). Similarly, we find qualitatively similar results for the treatment effects in the second period if we include only the period after the first experimental period (week 50 onwards).¹¹

Besides the average treatment effects, we are interested in how these treatment effects depend on the gender of the store manager, the gender composition of the store’s employees, the interaction between these two, and the store’s team size. To study these issues, we add interaction effects to equation (1). Let E_t^1 be a dummy that takes value one for all observations in the first experimental period (weeks 44 - 49) and, similarly, let E_t^2 be a dummy that takes value one in the second experimental period (weeks 71 - 76). The effect of variable X_s^p on the effect of our treatments is estimated by:

$$y_{st} = \alpha_s + \theta_t + \gamma B_{st}^1 + \eta X_s^1 B_{st}^1 + \kappa X_s^1 E_t^1 + \delta B_{st}^2 + \mu F_{st}^2 + \lambda X_s^2 B_{st}^2 + \pi X_s^2 F_{st}^2 + \psi X_s^2 E_t^2 + \nu R_{st} + \varepsilon_{st} \quad (2)$$

where X_s^p is the value of the variable X for store s in experimental period $p \in \{1, 2\}$ as taken from the personnel files received just before period p .¹² Again, we allow for differences between the estimated effect of variable X_s^p on the effect of the bonus treatment relative to

¹¹We cannot, however, identify any possible carry-over effects of first-period assignment to second-period treatment effects, as assignment in the second period was not completely random (see Figure 2). Note, however, the long time lag between the first and the second round (20 weeks).

¹²Obviously, in most of our regressions X_s^p is a vector of variables and, likewise, η , κ , λ , π , and ψ are vectors of coefficients. Note that, if we would have more frequent data from the personnel files, we could have taken up X_{st} as a separate regressor.

the feedback treatment between the first and second experimental period (η versus $\lambda - \pi$). The inclusion of the E_t^1 and E_t^2 terms is necessary to obtain difference-in-difference estimates of the effect of X_s^p on the treatment effects. Not including these terms would imply that period-specific effects of X_s^p on stores' sales growth would be erroneously picked up by the interaction effects with the treatments (i.e. by η , λ , and π).

4 Results

The first column in Table 3 gives the results of estimating (1). Focussing on the second period, we find that the bonus treatment and the feedback treatment both have a statistically significant positive effect on weekly sales growth. The size of the effects is also economically significant, as stores in the bonus and feedback treatment achieve 4.8 and 6.9 percentage points additional sales growth, respectively.¹³ The first column of Table 3 also shows that in the first period, the difference in the effects of the bonus and the feedback treatment on sales growth is small and statistically insignificant. The same holds in the second period: a Wald test on the restriction that the effect of the bonus and feedback treatments are equal has a P-value of 0.32. Hence, the financial reward in the bonus treatment did not trigger additional sales growth on top of the effect of the tournament that is also present in the feedback treatment.¹⁴

Next, we allow the treatment effects to vary by the gender of the store manager and by the gender composition of stores' employees. When we estimate (2) with a female manager dummy interacted with the treatments, we find only insignificant interaction effects (not reported for brevity). Hence, on average, the effects of the bonus and feedback treatments on sales growth do not differ between stores with a male manager and those with a female

¹³Lack of data on the absolute value of sales and profit margins implies that we cannot establish whether this increase in sales outweighs the cost of the tournaments. However, the company's management was content with these results.

¹⁴Furthermore, we cannot reject the hypothesis that $\gamma = \delta - \mu$ in (1), i.e. that the difference in the differences between the effects of the bonus and the feedback treatments between the first and second period are equal (p-value is 0.18).

manager. Similarly, we find insignificant interaction effects if we interact the treatments in (2) with the percentage of female employees in a store. Thus, across all stores, we find no evidence that the gender composition of teams influences the effects of competition. However, interacting the store managers' gender and the gender composition of store employees reveals an interesting pattern. The second column of Table 3 gives the results of estimating (2), where the treatments are interacted with both a female manager dummy and the percentage of female employees, as well as interacted with the *interaction* between the female manager dummy and the percentage of female employees. Thus, we allow for different effects of the gender composition of the stores' personnel on the effect of competition in stores with a male manager compared to stores with a female manager.¹⁵

Let us first focus on the second period. Remarkably, we find that for both treatments, the sign of the effect of the percentage of female employees in a store depends on the gender of the store manager. In stores with a male manager, the effect of competition on sales growth significantly decreases in the share of female employees. By contrast, in stores with a female manager, a higher percentage of female employees increases the effect of the treatments. This effect is also statistically significant: Wald tests reject the hypotheses that the percentage of female employees does not affect the effect of the bonus and feedback treatment in female-managed stores, with p-values of 0.06 and 0.10, respectively.

The estimated treatment effects for various manager/employee combinations are depicted in Figures 4 and 5. Figure 4 gives the point estimates and the 96% confidence interval of the effect of the bonus treatment in the second period (relative to the control group) for male-managed and female-managed stores separately. Figure 4 clearly shows that the bonus treatment has been most effective in raising sales growth in male-led stores with a relatively high percentage of male employees. Male-managed stores achieve significantly higher sales growth in the bonus treatment as long as the percentage of women employed does not exceed 80%. In female-managed stores, the pattern is reversed: the estimated impact of the bonus treatment strongly increases with the percentage of female employees. The treatment

¹⁵In the estimation reported in Table 3, the percentage of female employees is mean-centered.

effect is nowhere statistically significant at the 5%-level, but in all-female stores, the point estimate of 8.4 percentage points additional sales growth is statistically significant at the 8%-level. Comparing male-led and female-led stores, we find statistically significant differences in treatment effect for stores employing less than 75% female employees and for those employing only female employees.

The effects of the feedback treatment (relative to the control group), as depicted in Figure 5, are similar to those of the bonus treatment. For stores where the manager and a high percentage of employees have the same gender, we find substantial positive effects on sales growth, which are significant for male-managed and female-managed stores alike. For other stores, we find smaller, insignificant treatment effects. The treatment effects differ significantly between male-led and female-led stores for those employing less than 60% female employees and for those employing more than 90% female employees.

Comparing the bonus treatment and the feedback treatment, we find no significant differences in how the gender composition of stores affects performance under competition. Figure 6 and 7 visualize the point estimates and the 96% confidence interval of the effect of the bonus treatment relative to the feedback treatment in the first and second period, respectively, for male-managed and female-managed stores separately. These point estimates are nowhere significant, and the same holds for the difference between the estimates for male-managed and female-managed stores. Hence, regardless of the gender of the store manager and the gender composition of the team, the monetary reward in the bonus treatment did not significantly affect performance on top of the effect of competition evoked by the feedback treatment.¹⁶

Overall, our findings give a nuanced picture of gender differences in manager-led team performance under competition. It is not gender per se that affects performance under competition, but rather the match between the team's manager and the gender composition

¹⁶Comparing Figures 6 and 7, one observes differences between the two periods in how stores' gender composition affects the bonus treatment relative to the feedback treatment. However, Wald tests show that these differences are not statistically significant. For instance, the hypothesis that the difference in the effects between the bonus and the feedback treatment in all-female stores is equal in both periods is not rejected, with a p-value of 0.13.

of the team: competition positively affects performance when the manager and a sufficiently high percentage of employees have the same gender. This suggests that managers succeed in making the competition appeal to team members of their own sex, but less so to team members of the opposite sex.

Alternatively, our findings could be partly driven by the team nature of the incentive scheme. Success in the tournaments depends on team performance, so that limiting free-riding behavior is essential. It might be that male managers are better at reducing free-rider behavior in a team with many male employees, and likewise for female managers with female-dominated teams. To verify whether this interpretation has any relevance, Table 4 gives the results of including the number of employees interacted with the treatments to the regression reported in the second column of Table 3. Across our stores, there is quite some variation in the number of employees, ranging from a minimum of 5 to a maximum of 20 employees. Nevertheless, we find no evidence for free-riding behavior: for neither of the treatments, the effect of competition depends on the number of employees.¹⁷ Also, the inclusion of the number of employees hardly affects the estimates for the gender effects on performance under competition. Finally, interacting the number of employees with the gender composition of stores yields results that do not support the interpretation that free-rider behavior is reduced in stores where the manager and a large part of the store’s employees are of the same gender; details are available upon request.

We have checked the robustness of our findings by controlling for several other variables. First, the results reported in the second column of Table 3 are not affected if we control for managers’ tenure (i.e. when X_s^p in (2) includes managers’ tenure). Tenure itself negatively affects the effect of competition, although only significantly so in the feedback treatment.¹⁸ Similarly, controlling for the average age of employees in a store does not affect the results, and employees’ age itself does not influence the effects of the treatments. Lastly, none of our

¹⁷The same conclusion is drawn when the gender interaction terms are excluded from the estimation (so that X_s^p in (2) includes only the number of employees).

¹⁸An explanation could be that there is more room for improvement in stores with inexperienced managers. Conversely, managers with high tenure may be more averse to change.

results is affected qualitatively when weighing employees by their full-time equivalent.

5 Concluding remarks

We have studied how teams led by a professional manager respond to competitive incentives. Overall, we find strong effects, even when there is no monetary reward to winning the competition. Further, our results suggest that the gender of the manager and the gender composition of the team jointly affect performance under competition. Male-led teams are more responsive to competition when a larger fraction of the team members is male. By contrast, female-led teams respond more strongly to competition when the fraction of female members is larger. If generalizable, our results give rise to some optimism about future reductions in the stark gender inequalities in executive positions in business and government. As women have massively entered the labor market over the last decades, work teams have become more gender-diverse. Our results suggest that this should put female managers on a more equal footing in contests for promotions.

References

- [1] Auriol, Emmanuelle, and Régis Renault (2008), Status and incentives, *RAND Journal of Economics*, vol. 39(1), pp. 305-326.
- [2] Azmat, Ghazala, and Nagore Iriberry (2009), The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students, Universitat Pompeu Fabra Working Paper 1148.
- [3] Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2009), Team Incentives: Evidence from a Field Experiment, Mimeo, University College London.
- [4] Bertrand, Marianne, and Kevin F. Hallock (2001), The Gender Gap in Top Corporate Jobs, *Industrial and Labor Relations Review*, vol. 55(1), pp. 3-21.
- [5] Besley, Timothy, and Maitreesh Ghatak (2008), Status Incentives, *American Economic Review*, vol. 98(2), pp. 206-211.
- [6] Blanes i Vidal, Jordi, and Mareike Nossol (2009), Tournaments without Prizes: Evidence from Personnel Records, Mimeo, London School of Economics.
- [7] Booth, Alison L., and Patrick J. Nolen (2009), Choosing to Compete: How Different are Girls and Boys?, IZA Discussion Paper No. 4027.
- [8] Croson, Rachel, and Uri Gneezy (2008), Gender Differences in Preferences, *Journal of Economic Literature*, forthcoming.
- [9] Dargnies, Marie-Pierre (2009), Does Team Competition Eliminate the Gender Gap in Entry in Competitive Environments?, Centre d'Economie de la Sorbonne Working Paper 2009.06.
- [10] Datta Gupta, Nabanita, Anders Poulsen, and Marie-Claire Villeval (2005), Male and Female Competitive Behavior – Experimental Evidence, GATE Working Paper 05-12.

- [11] Dohmen, Thomas, and Armin Falk (2006), Performance Pay and Multi-dimensional Sorting: Productivity, Preferences and Gender, IZA Discussion Paper No. 2001.
- [12] Duchon, Dennis, Stephen G. Green, and Thomas D. Taber (1986), Vertical Dyad Linkage: A Longitudinal Assessment of Antecedents, Measures, and Consequences, *Journal of Applied Psychology*, 71(1), pp. 56-60.
- [13] Frey, Bruno S., and Susanne Neckermann (2008), Awards - A View from Psychological Economics, *Journal of Psychology*, vol. 216(4), pp. 198-208.
- [14] Gneezy, Uri, Kenneth L. Leonard, and John A. List (2008), Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society, *Econometrica*, forthcoming.
- [15] Gneezy, Uri, Muriel Niederle, and Aldo Rustichini (2003), Performance in Competitive Environments: Gender Differences, *Quarterly Journal of Economics*, vol. 118(3), pp. 1049-1074.
- [16] Gneezy, Uri, and Aldo Rustichini (2004), Gender and Competition at a Young Age, *American Economic Review*, vol. 94(2), pp. 377-381.
- [17] Green, Jerry R., and Nancy L. Stokey (1983), A Comparison of Tournaments and Contracts, *Journal of Political Economy*, vol. 91(3), pp. 349-364.
- [18] Harrison, Glenn, and John A. List (2004), Field Experiments, *Journal of Economic Literature*, vol. 42(4), pp. 1009-1055.
- [19] IPU (2008), *Women in Politics: 2008*, poster published by Inter-Parliamentary Union and United Nations Division for the Advancement of Women.
- [20] Jurajda, Stepan, and Daniel Münich (2008), Gender Gap in Admission Performance under Competitive Pressure, CERGE-EI Working Paper no. 371.

- [21] Lavy, Victor (2008), Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-Based Pay Tournaments among Teachers, NBER Working Paper 14338.
- [22] Lazear, Edward P., and Sherwin Rosen (1981), Rank-Order Tournaments as Optimum Labor Contracts, *Journal of Political Economy*, vol. 89(5), pp. 841-864.
- [23] Mai-Dalton, Renate R., and Jeremiah J. Sullivan (1981), The Effects of Manager's Sex on the Assignment to a Challenging or a Dull Task and Reasons for the Choice, *Academy of Management Journal*, vol. 24(3), pp. 603-612.
- [24] March, James G., and Herbert A. Simon (1958), *Organizations*, New York: John Wiley.
- [25] McNeilly, Kevin M., and Frederick A. Russ (2000), Does Relational Demography Matter in a Personal Selling Context?, *Journal of Personal Selling and Sales Management*, vol. 20(2), pp. 279-288.
- [26] Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007), Contests for Status, *Journal of Political Economy*, vol. 115(2), pp. 338-363.
- [27] Nalebuff, Barry J., and Joseph E. Stiglitz (1983), Prices and Incentives: Towards a General Theory of Compensation and Competition, *Bell Journal of Economics*, vol. 14(1), pp. 21-43.
- [28] Neckermann, Susanne, and Michael Kosfeld (2008), Working for Nothing? The Effect of Awards on Employee Performance, mimeo, University of Zurich.
- [29] Niederle, Muriel, and Lise Vesterlund (2007), Do Women Shy Away from Competition? Do Men Compete Too Much?, *Quarterly Journal of Economics*, vol. 122(3), pp. 1067-1101.
- [30] Örs, Evren, Frédéric Palomino, and Eloïc Peyrache (2008), Performance Gender-Gap: Does Competition Matter?, CEPR Discussion Paper No. 6891.

- [31] Paserman, Daniele M. (2007), Gender Differences in Performance in Competitive Environments: Evidence from Professional Tennis Players, mimeo, Hebrew University.
- [32] Pelled, Lisa Hope, and Katherine R. Xin (2000), Relational Demography and Relationship Quality in Two Cultures, *Organization Studies*, vol. 21(6), pp. 1077-1094.
- [33] Polachek, Solomon W. (1981), Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure, *Review of Economics and Statistics*, vol. 63(1), pp. 60-69.
- [34] Snizek, William E., and Cecily C. Neil (1992), Job Characteristics, Gender Stereotypes and Perceived Gender Discrimination in the Workplace, *Organization Studies*, vol. 13(3), pp. 403-427.
- [35] Tsui, Anne S., and Charles A. O'Reilly III (1989), Beyond Simple Demographic Effects: The Importance of Relational Demography in Superior-Subordinate Dyads, *Academy of Management Journal*, vol. 32(2), pp. 402-423.
- [36] Varma, Arup, and Linda K. Stroh (2001), The Impact of Same-Sex LMX Dyads on Performance Evaluations, *Human Resource Management*, vol. 40(4), pp. 309-320.
- [37] Wayne, Sandy, Robert C. Liden, and Raymond T. Sparrowe (1994), Developing Leader-Member Exchanges, *American Behavioral Scientist*, vol. 37(5), pp. 697-714.
- [38] Wirth, Linda (2001), *Breaking Through the Glass Ceiling: Women in Management*, Geneva: International Labour Office.

Tables and Figures

Figure 1: Weekly percentage sales growth compared to the same week the year before

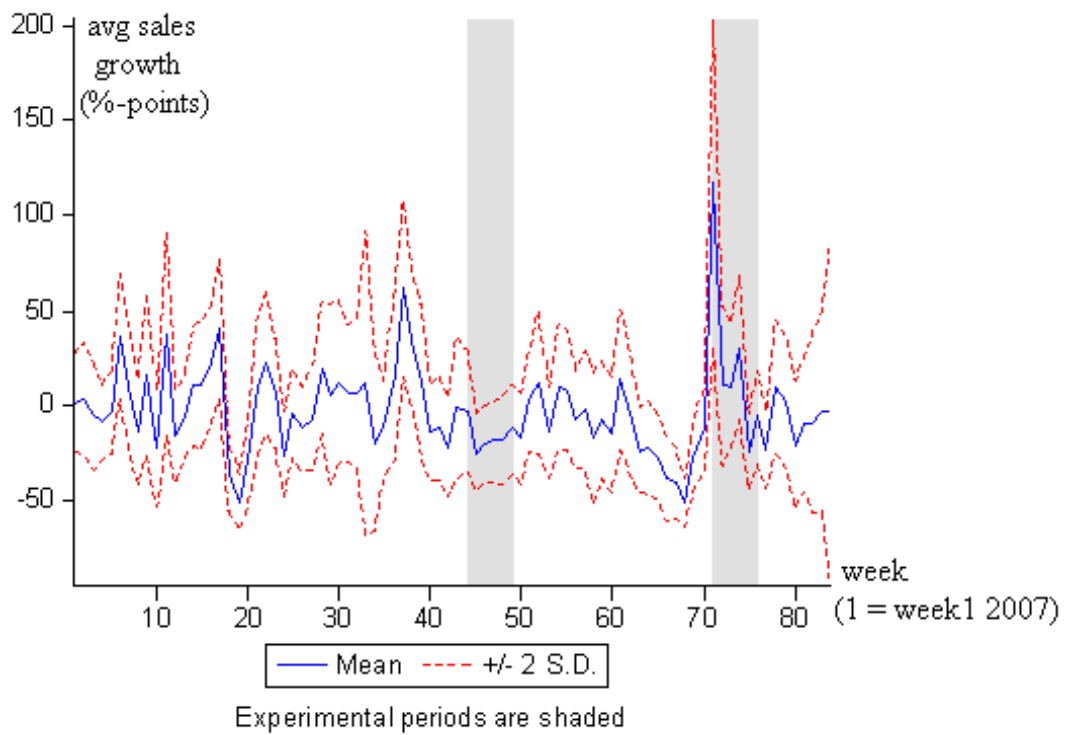


Figure 2: Overview of the experimental set-up and timing of events

week 1 - 43	week 44 - 49	week 50 - 70	week 71 - 76	week 77 - 84
personnel file	bonus treatment	updated personnel file	control	
	feedback treatment		feedback	
			bonus treatment	

Figure 3: Distribution of the percentage of female employees by store managers' gender

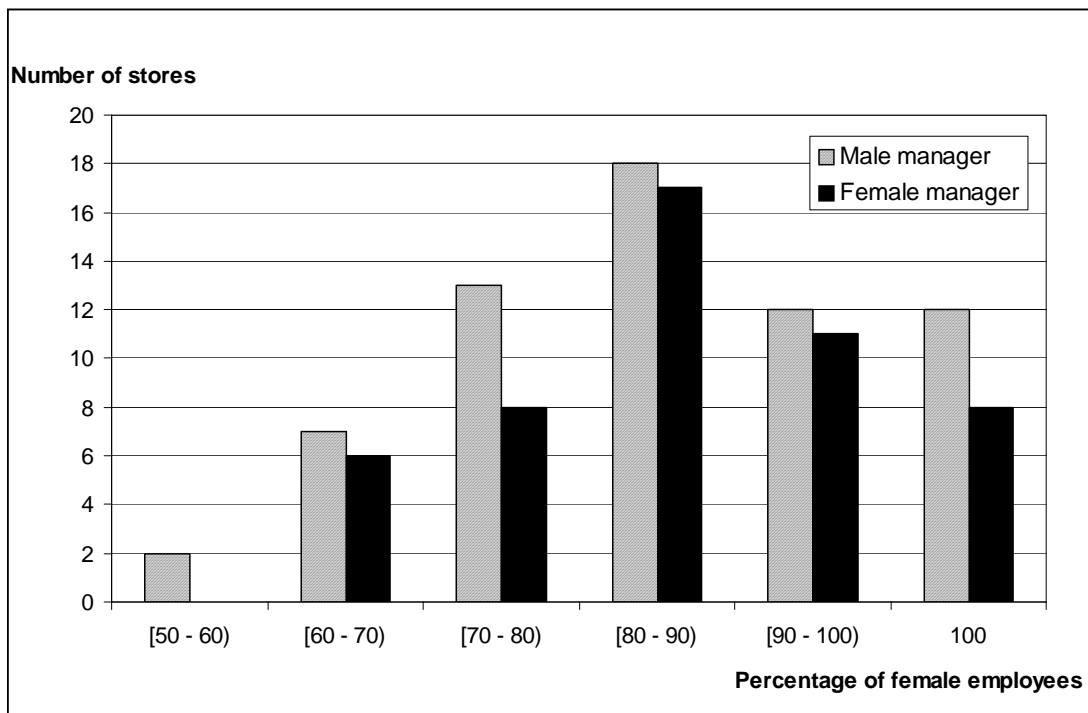


Table 1: Descriptive statistics

	Second period assignment							
	All stores		Bonus treatment		Feedback treatment		Control group	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Weekly sales growth (percentage points)	-2.44	29.87	-2.62	29.21	-2.22	30.90	-2.32	30.09
Female manager ^a	0.44		0.48		0.36		0.43	
Age manager ^a	38.7	9.3	39.2	10.3	38.2	7.6	38.1	9.1
Tenure manager ^a	11.9	8.7	12.9	9.3	11.7	8.7	10.7	7.8
Number of employees	11.8	3.2	11.8	3.4	11.2	2.7	12.2	3.5
Percentage of female employees	84.2	12.6	85.1	12.3	81.9	13.9	84.7	12.1
Average age of employees	25.2	4.0	26.0	4.5	24.7	3.3	24.4	3.3
Renovated store	0.13		0.17		0.07		0.13	
Number of stores	128		60		30		38	

Figures are the averages of the values in the two personnel files (see Figure 2).

F-tests show that none of the differences in means between the two treatment groups and the control group is statistically significant at the 10% level.

^a The first personnel file contained information on 122 store managers, and the second personnel file had information on 119 store managers; see footnote 10. The missing observations are not correlated with either of the treatments.

Table 2: Descriptive statistics by store managers' gender

	Gender manager			
	Male		Female	
	Mean	Std	Mean	Std
Weekly sales growth (percentage points)	-2.07	29.52	-2.46	30.79
Age	39.4	9.5	37.8	9.1
Tenure*	14.9	9.5	8.3	6.4
Number of employees*	12.6	3.2	10.7	3.0
Percentage of female employees	83.7	13.0	85.8	11.5
Average age of employees	26.0	3.6	24.6	3.7
Renovated store	0.17		0.10	
Number of stores	64		50	

This table includes only the 114 stores used in the analysis of the gender effects, see footnote 10.

Apart from sales growth, the figures are the averages of the values in the two personnel files.

* Difference in means for male and female managers statistically significant at the 5% level.

Table 3: The effect of competition on sales growth: gender differences

Independent variables	Dependent variable: Sales growth			
	(1)		(2)	
First period:				
Base category = Feedback treatment				
Bonus treatment	0.995	(0.828)	1.763	(1.257)
Female Manager x Bonus treatment			-1.590	(1.800)
Female Manager x Experimental period 1			-1.057	(1.263)
% Female employees x Bonus treatment			-0.058	(0.101)
% Female employees x Experimental period 1			0.168	(0.079)**
Female Manager x % Female employees x Bonus treatment			0.172	(0.163)
Female Manager x % Female employees x Experimental period 1			-0.187	(0.111)*
Second period:				
Base category = Control group				
Bonus treatment	4.796	(1.960)**	3.537	(2.394)
Feedback treatment	6.902	(2.387)***	3.346	(2.799)
Female Manager x Bonus treatment			-1.832	(3.928)
Female Manager x Feedback treatment			4.098	(4.941)
Female Manager x Experimental period 2			3.234	(3.056)
% Female employees x Bonus treatment			-0.384	(0.199)*
% Female employees x Feedback treatment			-0.416	(0.185)**
% Female employees x Experimental period 2			0.413	(0.136)***
Female Manager x % Female employees x Bonus treatment			0.831	(0.312)***
Female Manager x % Female employees x Feedback treatment			1.019	(0.414)**
Female Manager x % Female employees x Experimental period 2			-0.547	(0.230)**
Renovated store	8.315	(1.205)***	7.866	(1.268)***
Period fixed effects	YES		YES	
Store fixed effects	YES		YES	
Number of stores included	128		114	
Observations	10484		9326	
R ²	0.693		0.705	
Log likelihood	-44292.3		-39284.8	

***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

% Female employees is mean-centered.

Robust standard errors in parentheses.

Table 4: The effect of competition on sales growth: team size effects

		Dependent variable: Sales growth	
Independent variables			
First period:			
Base category = Feedback treatment			
Bonus treatment	2.010	(1.309)	
Female Manager x Bonus treatment	-1.768	(1.885)	
Female Manager x Exp. period 1	-0.307	(1.290)	
% Female employees x Bonus treatment	-0.137	(0.108)	
% Female employees x Exp. period 1	0.246	(0.088)***	
Female Manager x % Female employees x Bonus treatment	0.214	(0.164)	
Female Manager x % Female employees x Exp. period 1	-0.210	(0.113)*	
Number of employees x Bonus treatment	-0.215	(0.268)	
Number of employees x Exp. period 1	0.461	(0.189)**	
<hr/>			
Second period:			
Base category = Control group			
Bonus treatment	3.640	(2.650)	
Feedback treatment	3.838	(2.978)	
Female Manager x Bonus treatment	-1.963	(4.211)	
Female Manager x Feedback treatment	1.973	(5.320)	
Female Manager x Exp. period 2	3.871	(3.348)	
% Female employees x Bonus treatment	-0.355	(0.202)*	
% Female employees x Feedback treatment	-0.417	(0.185)**	
% Female employees x Exp. period 2	0.419	(0.137)**	
Female Manager x % Female employees x Bonus treatment	0.812	(0.311)***	
Female Manager x % Female employees x Feedback treatment	1.041	(0.420)**	
Female Manager x % Female employees x Exp. period 2	-0.531	(0.229)**	
Number of employees x Bonus treatment	0.034	(0.588)	
Number of employees x Feedback treatment	-1.051	(1.062)	
Number of employees x Exp. period 2	-0.268	(0.421)	
Renovated store	7.884	(1.270)***	
Period fixed effects	YES		
Store fixed effects	YES		
<hr/>			
Number of stores included	114		
Observations	9326		
R ²	0.705		
Log likelihood	-39282.0		

***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

% Female employees and Number of employees are mean-centered.

Robust standard errors in parentheses.

Figure 4: Gender differences in the effect of the bonus treatment

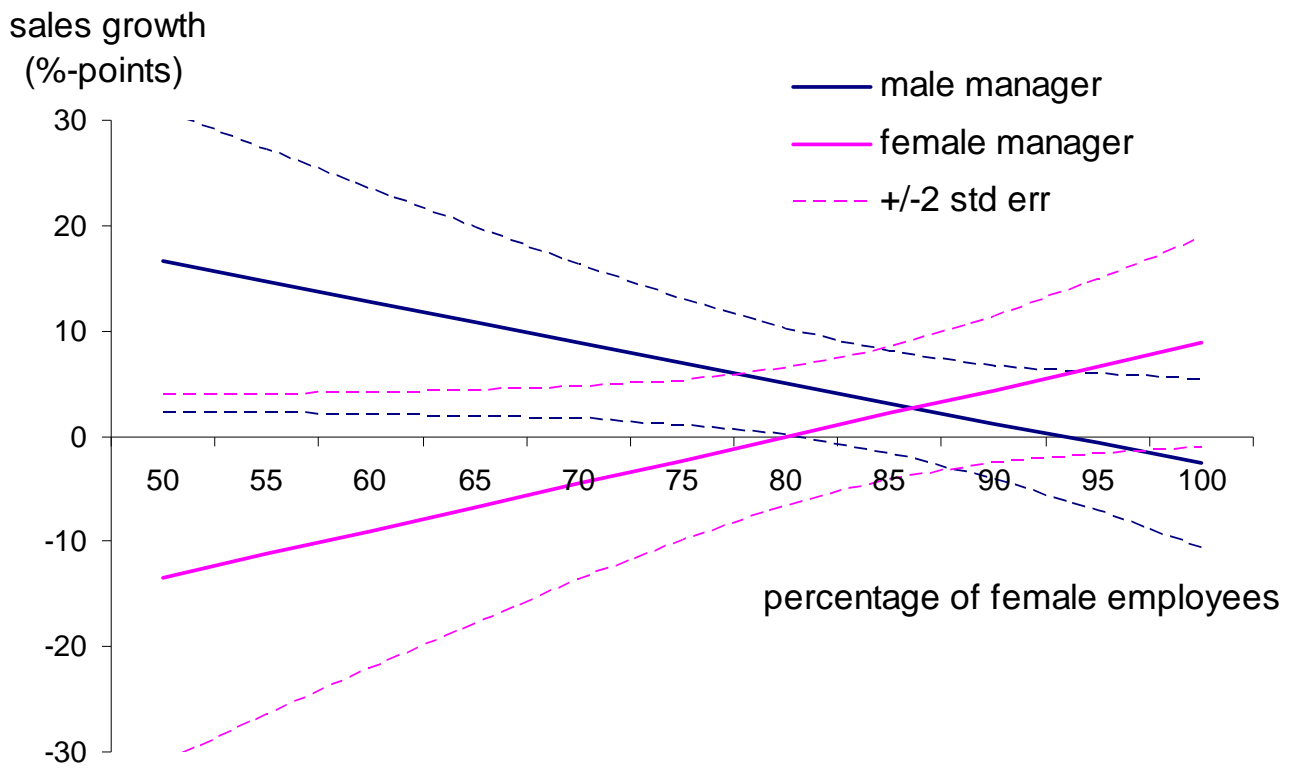


Figure 5: Gender differences in the effect of the feedback treatment

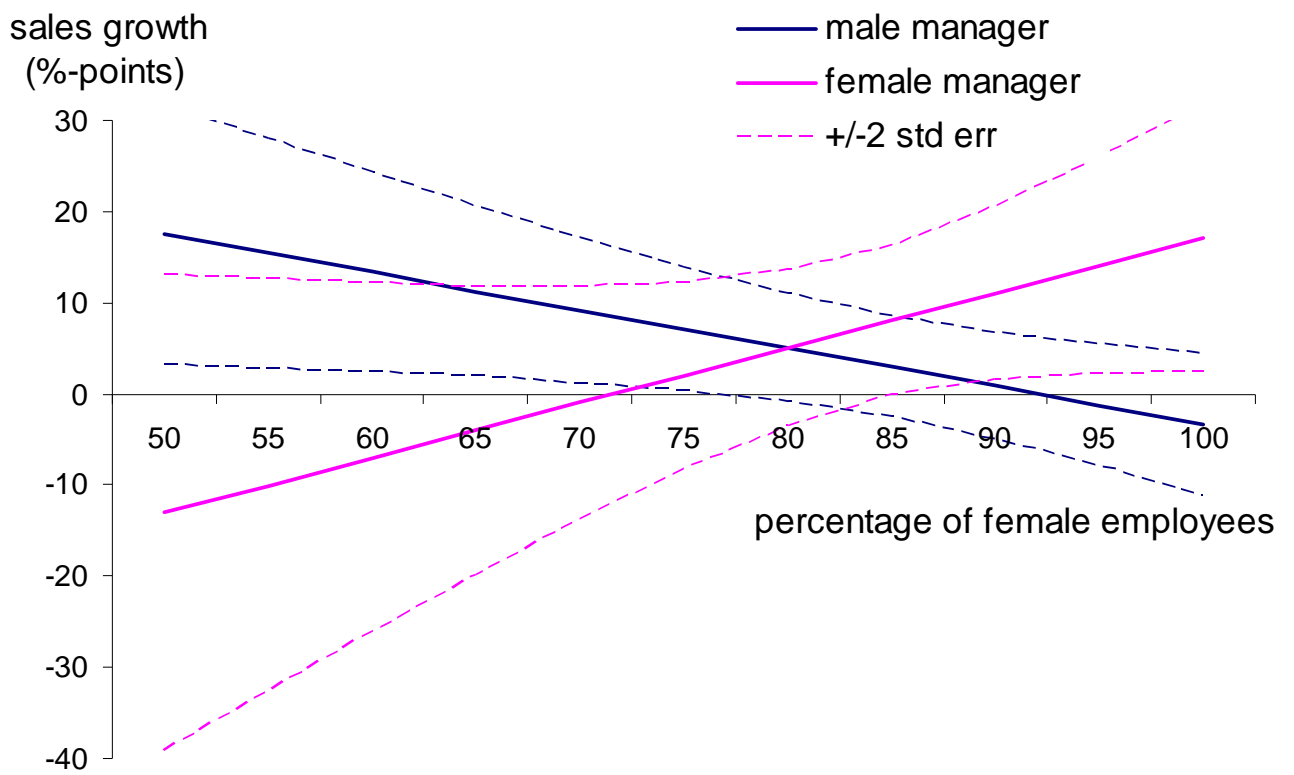


Figure 6: Gender differences in the effect of the bonus treatment (relative to feedback, first period)

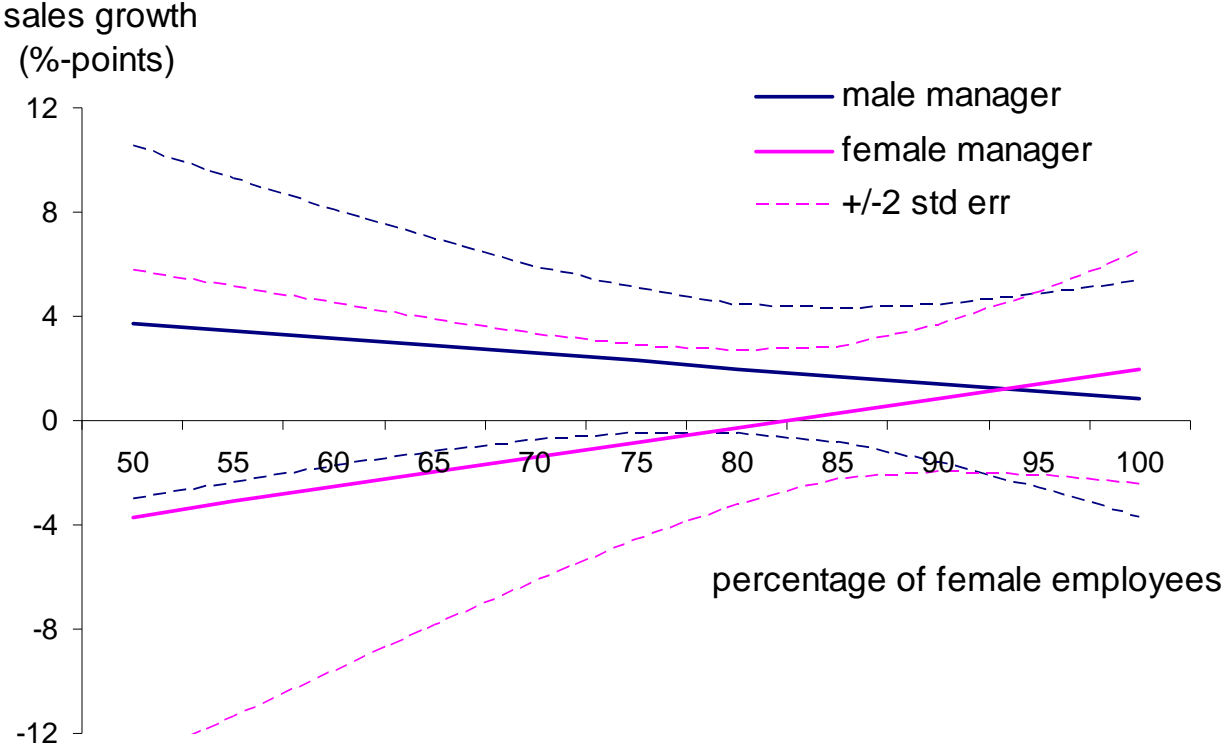


Figure 7: Gender differences in the effect of the bonus treatment (relative to feedback, second period)

