Can Promotion Tournaments Produce Bad Managers? Evidence of the "Peter Principle"

Alan Benson, Danielle Li, and Kelly Shue*

November 7, 2016

Abstract

The best worker is not always the best candidate for manager. In these cases, do firms promote the best potential manager or the best worker in their current job? Using microdata on the performance of sales workers at 214 firms, we find evidence consistent with the Peter Principle: firms prioritize current job performance when making promotion decisions at the expense of other observable characteristics that better predict managerial performance. We estimate that the costs of managerial mismatch are substantial, suggesting that firms make inefficient promotion decisions or that the incentive benefits of emphasizing current performance is also high.

*Respective affiliations are: University of Minnesota Carlson School of Management; Harvard Business School; and University of Chicago Booth School of Business and NBER. We thank Steve Kaplan, Eddie Lazear, Michaela Pagel, Brian Phelan, Thomas Peeters, Lamar Pierce, Felipe Severino, and participants at the FOM Conference, Harvard Business School, MIT Sloan School of Management, Carlson School of Management, Midwest Economics Association Meetings, NBER Summer Institute, Wharton People Analytics Conference, and Yale Junior Finance Conference for helpful comments. We thank Menaka Hamplole and Leland Bybee for excellent research assistance. Correspondence: bensona@umn.edu.

The best workers may always make the best managers being a manager often requires a different set of skills. When this is the case, do firms promote someone who excels in her current position, or someone who is likely to excel as a manager? If firms promote based on current performance, then firms may end up with worse managers. Yet if firms promote based on the traits that make good managers, then firms may pass over higher performing workers, forfeiting the power of promotions to encourage workers to perform well in their current roles.

Using detailed microdata on sales workers in US firms, we provide the first large scale evidence that firms prioritize current performance in promotion decisions at the expense of promoting the best potential managers. Our findings lend support for the "Peter Principle," which, in its extreme form, states that firms promote competent workers until they become incompetent managers (Peter and Hull 1969). Sales is an attractive setting to study the Peter Principle since direct sales skills and sales management skills sharply differ (Fairburn and Malcomson 2001; Waldman 2003).[1] Unlike other potential settings for the Peter Principle (such as promoting excellent engineers into management positions), the performance of sales workers is readily measurable. Sales is also an important and policy-relevant setting, as sales and sales-related occupations account for 10.5 percent of the US labor force (Bureau of Labor Studies 2015).

Our analysis uses new data that are uniquely well suited to examine whether firms' promotion policies maximize managerial match quality. These data, provided by a firm that offers sales performance management software over the cloud, include standardized measures of sales transactions and organizational hierarchy for a panel of 48,209 workers matched to 5,369 managers at 214 different US-based client firms in a range of industries from 2005 to 2011. The transaction-level data include both the value of sales credited and to whom credit for each sale is split. For sales workers, we use employment history and crediting data to examine promotion as a function of sales performance (the dollar value of sales), sales collaboration (the number of other salespeople with whom they shared credit on transactions), and other observable worker characteristics. For promoted managers, we calculate their "manager value added" in shaping their subordinates' sales

[1]Baker, Jensen and Murphy (1988) state that "in many cases, the best performer at one level in the hierarchy is not the best candidate for the job one level up—the best salesman is rarely the best manager." Deutsch (1986) points out that "American companies have always wrestled with ways to keep the Peter Principle at bay–to prevent competent salesmen, for example, from rising to become incompetent sales managers."

performance, i.e., each manager's contribution to improving their subordinates' sales, controlling for other potentially confounding factors such as fixed effects for each subordinate and fixed effects for each firm-year-month.

We begin by showing that past sales performance is a strong positive predictor of promotion in managerial positions. A doubling in sales credits increases the probability that a worker is promoted by sixteen percent relative to the base probability of promotion. Of course, promotion policies that favor strong sales performance need not be evidence of the Peter Principle if past sales performance also strongly predicts managerial performance. However, among observed promotions, we find that pre-promotion sales performance negatively predicts manager value added: a doubling in the new manager's pre-promotion sales corresponds to a 10 percent decline in the performance of each of the newly-promoted manager's subordinates. Equivalently, newly-promoted managers who were relatively poor salespeople prior to their promotions tend to see their subordinates' performance subsequently improve. This negative correlation is consistent with the Beckerian insight about selection and discrimination: if firms' promotion policies "discriminate" against poor sales performers, then poor sales performers who are nevertheless promoted should be better managers.

Our analysis also identifies another observable worker characteristic that firms appear to underweight in promotion decisions: sales collaboration experience. Pre-promotion collaboration experience positively predict managerial performance. However, workers with more collaboration experience are not significantly more likely to be promoted after accounting for firm-level time trends.

Next, we show that these findings are robust to potential sample selection concerns. We face a data limitation in that we do not observe latent managerial potential for all workers. Instead, we observe managerial performance for the subset of workers that are promoted. Therefore, we began by measuring the relation between pre-promotion worker characteristics and managerial performance for the sample of promoted workers. Because promotions are not random, selection may bias our estimates of the predictive relationship between sales and managerial quality. For example, suppose that sales and a second unobserved quality both positively predicts managerial quality, and that firms correctly value both qualities when making promotion decisions. In this

case, promoted high sales workers may be weaker on unobservables than promoted low sales workers because their sales compensate for other managerial weaknesses. This type of selection would negatively bias our estimates of the true relationship between sales and managerial quality; for sufficiently negative biases, it may even lead us to estimate a negative coefficient when the true relationship is positive.

To address these types of problems, we apply a model of promotions based on the Heckman (1976) selection model. We recover the average relation between sales and latent managerial performance for the full sample of workers, using variation in promotion thresholds that are unrelated to managerial performance. In particular, we instrument for promotion, i.e., selection into the sample of promoted managers, using the average firm- and industry-level promotion rates within a time period. These average promotion rates reflect time-varying vacancies and firm- and industry-level demand for managers and are strongly positively correlated with a worker's probability of promotion. Of course, high average promotion rates may also reflect strong consumer demand, which will positively affect the performance of all sales workers and may thus be correlated with managerial performance. In our setting, however, we measure a manager's quality as his or her value-added to subordinate sales, net of worker and firm-year-month fixed effects. Because our instruments vary only at the company-month level, they are by construction orthogonal to our measures of manager value added. After accounting for selection, we continue to find a strong negative relation between pre-promotion sales performance and manager value added. We also provide additional evidence in Section 4 that our results are not driven by mean reversion, non-linear relationships, or non-random matching of subordinates to managers.

Our main empirical findings provide support for the Peter Principle: firms promote based on current job performance even though pre-promotion sales negatively predicts managerial performance and other observable characteristics positively predict managerial performance. Because our approach allows us to recover estimates of predicted manager value added for the full sample of workers, we can also estimate productivity losses associated with firms' promotion policies by comparing the predicted managerial performance of actual promoted managers to the predicted performance under alternative promotion rules. We estimate that firms can improve the sales performance of

subordinates by up to 30 percent if they follow an alternative promotion policy that maximizes managerial performance.

We caution that our results do not necessarily imply that firms use suboptimal promotion policies or have mistaken beliefs. Promotion policies that favor strong sales performance may provide a variety of incentive benefits that justify the costs of managerial mismatch. For example, promoting based on current job performance may help preserve tournament incentives (Lazear and Rosen 1981). Prioritizing objective performance measures in promotions may also improve incentives by avoiding favoritism (Prendergast 1998) and maintaining fairness norms. Promotion policies based on verifiable performance metrics such as sales may also discourage the manipulation of other, more fungible performance metrics such as credit sharing and collaboration experience (DeVaro and Gürtler 2015).[2] What our results do show is that the costs of not promoting the best potential managers are high: our counterfactual estimates suggest that firms are willing to forgo up to a 30% improvement in subordinate performance to achieve better incentives.

Finally, our findings raise the question of whether firms can reduce the costs associated with the Peter Principle by using alternative incentive schemes. Strong pay for performance may offset the need to reward current job performance with promotions. To examine this, we explore how promotion policies vary with pay-for-performance at the firm-year level. An advantage of our data is that we observe commission and bonus compensation for a subset of workers across a large number of firms. For this subset of workers and firms, we find that firms with greater pay for performance—defined as those in which commissions and bonuses are larger relative to fixed compensation—have promotion policies that are less sensitive to worker sales performance.

These findings are consistent with the idea that pay-for-performance incentives can offset the need to provide incentives through promotion tournaments. However, our results do not imply that pay-for-performance can eliminate the costs associated with the Peter Principle. Rather, pay-for-performance may be an expensive substitute for tournament incentives from the point of view of firm management and shareholders. Further, relative to other occupations, sales is associated with very high pay-for-performance on average, and yet we continue to find evidence of

[2]In our setting, workers could potentially add fake collaborators by sharing credits if firms began to heavily weight collaboration experience in promotion decisions.

the Peter Principle within the sales occupation overall.

This study offers the first empirical evidence of the magnitude of the Peter Principle using data on promotions across a large number of firms. Although theoretical work and reviews have hypothesized that tournament incentives may yield mismatch in promotions (Fairburn and Malcomson 2001; Waldman 2003; Lazear 2004), scarce empirical research has tested the Peter Principle directly. Our work is most closely related to Grabner and Moers (2013), which shows that a bank places less weight on current job performance when a promotion would be to a job performing dissimilar tasks. However, Grabner and Moers uses data from a single firm and does not attempt to estimate the cost of the Peter Principle. Understanding the costs associated with the Peter Principle is important because it helps explain a variety of organizational practices, such as the use of dual career tracks for individual contributors and managers, or the use of separate criteria for making bonus and promotion decisions.

Our findings also relate to the literature exploring the declining popularity of internal promotions and rising popularity of external managers or directors (e.g., Murphy and Zabojnik 2004). The decision to hire an external manager must weigh the benefits of expanding the field of candidates to improve the quality of an eventual match against the costs of reducing incentives for internal candidates. Our research is also motivated by recent findings in Kaplan, Klebanov and Sorensen (2012) and Kaplan and Sorensen (2016), which show that general skills, execution ability, and interpersonal skills, as measured on executive personality assessment exams, are associated with executive career progression and performance. These findings underscore the possibility that hiring managers based on performance in lower level job roles, rather than potential managerial ability, can be extremely costly.

1 Setting and data

Our data come from a firm that offers sales performance management (SPM) software over the cloud. The firm's clients input their employee records, organizational hierarchies, and sales transactions into the software, which then calculates pay for each individual worker. Transaction inputs can be entered manually or linked to order management and customer relationship management

(CRM) software. Pay outputs are typically linked directly to payroll software. The software also provides reporting and analysis. Sales workers and sales managers can log in to the SPM software's website to view their sales credits, progress toward quotas, commissions, and other data. The software can also generate reports for use in auditing and compliance with Sarbanes-Oxley.

[Table 1]

The data include 214 client firms, 48,209 sales workers and 5,369 sales managers. Client firms' most-represented industries include manufacturing (62 firms), information (56 firms), and professional services (38 firms). Sales workers in our data sample tend to be more highly skilled and compensated than the typical US sales worker. In 2011, sales occupations employed 13.6 million workers at a median wage of $24,840 (U.S. Census Bureau 2011), about half of whom worked in retail sales. By contrast, sales workers in our data predominantly work in business-to-business sales and earn a median monthly commission of $3,584 ($43,008 annually), not including base salary and bonuses. Table 1 provides descriptive statistics for sample coverage. All firms have at least one complete fiscal year of data, and no one firm constitutes more than 8% of person-months.

1.1 Overview of sales positions

Sales workers are typically assigned a market consisting of a territory, and a set of products or a type of client. Within their market, they are typically responsible for generating leads on potential new clients, making first contact, executing the initial sale, cross-selling other products, selling upgrades, and maintaining relationships. The sales industry refers to this process as the sales cycle.

The primary measure of a salesperson's performance is the total dollar value of the sales that he or she makes or contributes to. Our data include 156 million sales transactions tied to individual workers. Table 1 describes the distribution of sales generated. Because sales tend to be intermittent and can vary over the year, we report rolling averages of sales credits in the previous 12 months. The quartiles for monthly worker sales are $42,883, $272,987, and $1.50 million. Reflecting the wide and skewed distribution of sales standards across markets in which workers operate, the mean of this figure is $3.57 million.

[Figure 1]

Figure 1 also illustrates the skewness in the distribution of sales. The top left panel presents a histogram for the raw distribution of worker-level monthly sales (measured as 12 month rolling averages). The middle left panel plots the log of monthly sales, and shows that this follows a less skewed distribution. The bottom left panel, which reflects our measure of sales performance, shows the residual distribution of monthly sales after controlling for firm-year-month fixed effects. In other words, we measure sales performance as the recent performance of a sales worker compared to others in their same company at that time. Even with these fixed effects, we still observe wide variation in sales credits across workers. The interquartile range of residual log sales credits is 1.99. This means that, among rank and file sales workers in the same firm in the same month, a worker in the 75th percentile generates $e^{1.99} = 7.32$ times as much revenue as a 25th percentile worker. Although this difference is stark, it's also consistent with the so-called "80-20 rule," a well-known adage in the sales industry that states that the top 20% of the sales force is responsible for 80% of sales; in other words, sales follow a Pareto distribution.

In addition to total sales, we also obtain data on collaboration experience. In the complex business-to-business sales settings that constitute the majority of our data, sales transactions are often credited to more than one worker. For example, a relationship manager may be a client's single point of contact. For specialized products and services, the relationship manager may consult a product specialist, and, if a sale is made, both the relationship manager and the product specialist would receive a credit. For complex products and services, a single transaction can involve salespeople across many sales functions, products, and geographies. In our data, we observe all workers credited on a transaction and define a salesperson's collaboration experience as the number of distinct other workers with whom she shares credit on her transactions.

Table 1 provides summary statistics for collaboration and Figure 1 presents histograms of the distribution of the number of distinct collaborators each month, again measured as 12 month moving averages. Almost 50% of workers work alone, while the remainder vary greatly in their number of collaborators. This difference is not merely reflective of differences in work organization across firms or over time. The bottom-right panel of Figure 1 shows that even within the same

firm, in the same month, there is substantial variation in the extent to which workers collaborate on sales. The within firm-year-month interquartile range of sales collaborators is 0.90, signifying that the 75th percentile worker has $e^{0.90} = 2.45$ times as many collaborators as the 25th percentile worker.

This variation highlights two classic types of sales workers: those that are the only person credited on transactions and those that share sales credits on transactions. Indeed, much of the practitioner literature emphasizes different performance management practices for these groups. "Lone wolves" might be recruited for their self-confidence, resilience, and autonomy, and are stereotypically marked by their reluctance to share leads, best practices, and client relationship responsibilities with others in the organization. The most effective team players, by contrast, enable those around them by forwarding leads, crafting sales that include many others' territories and products, forwarding established clients to account managers, and developing team members so they can be effective in these capacities. These lead generation and origination activities would also generally entitle that salesperson to a portion of the sales made by others.

Within a firm-year-month, the correlation between our sales and collaboration measures is 0.28. While positive and statistically significant, the moderate correlation shows that there is substantial variation across these measures.

Table 1 also provides summary statistics for worker pay. The quartiles for monthly commission pay are \$2,509, \$3,584, and \$16,524. In addition, most workers also receive a fixed salary. Because our data provider's software is not designed to track or distribute base salaries, and instead treats base salaries as an optional field, base pay data can be missing or measured with error (for example, if it's not updated or the pay periods are unclear). Based on this limited data, however, we believe that the median worker in our sample receives between three and five thousand dollars per month in base pay.[3]

Our analysis uses monthly sales as the measure of pre-promotion sales performance, which has the advantage of being highly standardized, and after controlling for firm and time effects, has

[3]The salary figures reported in Table 1 are based on data that are unclear in units. We believe that the majority of salaries are reported in monthly terms but some may be reported annually, leading us to believe that true monthly salary is lower than the median \$6,911 figure reported in this table.

an easy interpretation. A limitation of our sales performance measure is that we do not observe the profit margins associated with sales transactions, which in principle are what firms should be maximizing. Nevertheless, we believe that the relative levels of sales credits among workers in the same firm and time offers a reasonable approximation of relative sales performance, and in the data, these sales mechanically determine the ultimate commissions. In theory, we could instead use worker compensation as a measure of sales performance. However, this approach would also have disadvantages. First, compensation can be difficult to interpret because firms can "pay for performance" by paying a high base rate and setting high standards for retention. Second, compensation doesn't always correspond to recent performance; for example, firms can pay workers origination commissions for renewals on sales that were made far in the past. Third, the base pay data can be unreliable since it's not required by the software and not directly linked to payroll. Therefore, we prefer relative sales credits as the primary measure of sales performance.

Note also that the worker tenure variable is censored on the left-hand side by the date the firm began using the SPM software. Therefore, we control for tenure within the SPM system and its interaction with whether tenure is potentially censored.

1.2 Overview of managerial positions

In our data, we observe the hierarchical structure linking managers to subordinates. We define a sales manager to be someone who has at least one subordinate. Managers typically have titles such as "territory manager," "sales director," "regional director," "regional manager," "sales engineer manager," and "regional vice president." The last panel of Table 1 summarizes the characteristics of these managers. On average, each manager has five subordinates. Managers have significantly higher commission and salary pay than workers on average, although the top salespeople still make more than the median managers (a potential selection issue relating to some high sales workers preferring not to be promoted to managerial positions is discussed in later sections).

Managers also have substantially different tasks. While sales workers are primarily engaged in direct sales activities, sales managers are responsible for building a high-performing sales team. A survey of first-line sales managers by the Sales Management Association reports that sales managers

spend the most time on performance management, followed by company administration, sales planning, selling and market development, and staff deployment. Performance management requires leadership, coaching, and training skills that may be imperfectly related to those used in direct sales activities. Administrative duties require general management knowledge so that the sales manager can interface with other functions, such as marketing and operations. Sales planning requires data analysis skills so that managers can read market research, set quotas, assign territories, monitor performance, and prioritize sales activities. Sales managers also oversee the development of playbooks that compile best practices and outline the company's strategy for selling their products. Successfully executing these activities reflects in the performance of their teams. For example, if the manager misreads market research, manpower could be misallocated to unproductive products or territories, quotas could be set at unattainably demotivating thresholds, or training could encourage salespeople to emphasize the wrong product features for their market.

Because sales managers are ultimately responsible for improving the performance of their subordinates, we measure managerial performance by examining how managers impact the sales performance of their subordinates. In general, any measure of managerial performance that relies on subordinate performance can be biased by non-random assignment of managers to subordinates. For example, if a manager is assigned to better-performing subordinates, then we may mistakenly credit the manager for the subordinates' higher sales numbers. Indeed, Table 8, presented in later sections, shows that managers with higher pre-promotion sales tend to be assigned to subordinates who had relatively high sales under their previous managers.

To address these concerns, we follow Lazear, Shaw and Stanton (2016), as well as a large literature on employer-employee, and teacher-student matched data (e.g., Abowd et al. 2001), and estimate a manager's performance as his or her *value-added*, controlling for each worker's average performance across different managers and firm-year-month time trends. Specifically, we extract the value-added measure from a regression of the form:

$$Log(1 + \text{Sales}_{imft}) \quad = \quad a + \delta_i + \delta_m + \delta_{f\times t} + e_{imft} \tag{1}$$

Here, the dependent variable is the log of worker i's sales performance under manager m in firm

f in year-month t. δ_i, δ_m, and $\delta_{f\times t}$ represent worker, manager, and firm-year-month fixed effects, respectively.[4] The coefficients of interest are those on the manager fixed effects, δ_m, which is the time-invariant component of a manager's quality or average value added.

The inclusion of both manager and worker fixed effects means that manager value added is identified from workers whom we observe under multiple managers. A manager's fixed effect is estimated as the average change in sales performance across all workers who switch to or from that manager. As such, a manager with high value added is one under whom workers perform above their individual mean across all the managers under whom they have worked. Whether a manager is assigned to strong or weak subordinates should not impact our measure of value added because a manager is only credited for improving the performance of his or her assigned subordinates. Further, macroeconomic and other firm-time specific conditions that may impact subordinate sales performance are also netted out through the firm-year-month fixed effects.

Our estimates may still suffer from bias if managers are non-randomly assigned to subordinates on the basis of the subordinates' time trends in performance. If some managers are systematically more likely to be assigned to subordinates whose sales performance is on an increasing trend, then Equation (1) may mistakenly attribute their improvement to the manager's ability. Potential systematic biases in how manager value added is estimated would only impact our results if the time trends in subordinate performance are further correlated with a manager's pre-promotion characteristics. In Section 4, we will show that this does not appear to be true in our data sample.

Another important factor is noise. In Equation (1), we use monthly sales as the outcome variable even though this measure is quite variable, with many workers making zero sales in some months and large sales in others.[5] Assuming workers with different time trends in monthly sales are not systematically more likely to be assigned to certain types of managers than others (something we test in Section 4), each manager's estimated fixed effect is an unbiased, albeit noisy, estimate of her managerial performance. In other words, each estimated manager fixed effect is equal to the

[4] We estimate this regression using the Stata package `felsdvreg`. Rather than estimating $\delta_{f\times t}$ directly, we demean the outcome variable by firm-year-month prior to estimation in order to reduce the computational demands of the regression specification.

[5] Because our estimates are based on worker movements across managers, we cannot use backward rolling averages because this would be contaminated by performance under previous managers.

manager's true quality plus noise. Our analysis will regress these manager fixed effects on variables representing the managers' pre-promotion characteristics. Error in the dependent variable in these regressions should not bias our estimates how *mean* differences in manager value added relate to pre-promotion characteristics. However, this additional noise does lead to an upward bias in our estimates of the *variance* of manager fixed effects.

[Table 2]

Table 2 describes our manager sample and the identifying variation we use to estimate manager value added. We observe 5,369 managers in our data, of whom we are able to estimate fixed effects for 2,231. This lowered number comes from the high bar required to identify manager fixed effects: in order to estimate a fixed effect for a given manager, we must observe that manager supervising multiple subordinates whose own fixed effects are known through their work under other managers. Many managers in our sample do not have subordinates who are observed under other managers and this is the main reason why we are only able to estimate manager fixed effects for approximately half of our manager sample.

Moreover, these manager fixed effects are estimated within groups of workers and managers who are connected through moves. For instance, a connected group might contain a manager, her new subordinates, the previous managers of those subordinates, and the other subordinates of those managers. Fixed effects for managers within the same connected group are comparable relative to a group-specific normalization. For the median firm in our sample, 96.5 percent of workers are part of this largest connected group. To make these fixed effects more comparable across firms, we further demean them by firm specific averages. Because manager fixed effects are estimated with varying precision, we weight summary statistics and regressions involving these fixed effects by the inverse variance of our estimates.

Finally, we do not observe pre-promotion performance for most managers with estimated fixed effects. This is because many of the managers were always managers in our sample. We have information on both manager value added and pre-promotion performance for 696 managers that are promoted during our sample period.

[Figure 2]

Figure 2 plots the distribution of our estimates of manager value-added. By construction, this has been demeaned to have a mean of zero. The 25th percentile of this distribution is -0.87, meaning that when assigned to a 25th percentile manager, worker's output is only $e^{-1.01} = 0.36$ of what it would have been under the mean manager. Conversely, when assigned to a 75th percentile manager, worker's output increases by a factor of $e^{0.99} = 2.69$. Note that this interquartile range may be large because it reflects real differences in managerial ability or because of noise in the estimation of manager fixed effects, which exaggerates the variance.[6]

2 What predicts promotion?

Our first empirical exercise examines how sales and collaboration experience predict promotion of front-line sales workers to managers:

$$\text{Promote}_{ift} = a_1\text{Sales}_{ift} + a_2\text{Collaborators}_{ift} + W_{ift} + \delta_{f\times t} + e_{ift} \tag{2}$$

We estimate Equation (2) on a worker-year-month level panel for worker i at firm f who have not yet been promoted as of year-month t. The dependent variable, Promote_{ift} is an indicator for whether a worker is promoted in the next month. Sales_{ift} is the log of one plus worker i's monthly sales credits, averaged over the past 12 months or for the worker's total tenure if tenure is less than 12 months. $\text{Collaborators}_{ift}$ is defined as the log of the unique number of sales collaborators in a month who have shared credits with worker i, again averaged over the past 12 months or for the worker's total tenure if tenure is less than 12 months. The other covariates W_{ift} include fixed effects for seven bins of a worker's team size. Team size is distinct from our main measure of collaboration experience. Team size represents the number of other sales workers who share the same manager (team members do not necessarily collaborate on sales transactions). W_{ift} also includes fixed effects for seven bins of worker tenure, interacted with whether tenure may be censored in the data. In

[6] As discussed earlier, regressions presented in later sections that estimate the mean expected difference in manager value added across managers with different pre-promotion attributes remain unbiased because noise in the dependent variable should not bias regression coefficients.

some specifications, we also control for the industry-wide and firm-wide promotion rates in the current month. Finally, we can control for firm-year-month fixed effects $\delta_{f\times t}$ which subsumes the industry or firm-wide promotion rates, which only vary at the firm-year-month level.

[Table 3]

Table 3 reports the regression results. Firms are significantly more likely to promote higher performing salespeople, and this result is robust across specifications. Controlling for firm-year-month fixed effects, Column 4 shows that a doubling in sales credits increases the probability that a worker is promoted in each month by 0.033 percentage points. This represents a sixteen percent increase relative to the base rate of being promoted in each year-month of 0.21 percent. We also find that collaboration experience positively predicts promotion. However, this correlation becomes insignificant once we control for the firm-level promotion rates or firm-year-month fixed effects.

[Table 4]

We find similar patterns if we estimate promotion probability using a probit model in Table 4 instead of a linear probability model. The estimates presented in Columns 3 and 4 of Table 4 will also be used in later analysis as the first stage of a Heckman selection model. In later analysis, we also examine non-linearities in the predictive relation between sales and promotion. We find in the first two columns of Table 9 that promotion probability is low and insensitive to sales for workers with very low sales, and then increasing with sales for workers with moderate or high levels of sales. Similarly, in Figures **??** and **??**, we document monotonic relationships between pre-promotion sales and collaborations, and, respectively, promotion propensities and manager value added.

[Figure 3]

These patterns are underscored by Figure 3, which plots the distribution of workers' fitted propensities to be promoted for different subsamples. For the top panel, we compute residualized sales, equal to the residual from a regression of log sales on the log number of collaborators and all other control variables in Column 4 of Table 3. We then plot the promotion propensity separately

for workers with residualized sales above or below the median in the sample. A worker with high residualized sales is in the top half of sales in her firm-year-month, among workers who have the same number of collaborators. Our results show that higher performing salespeople are much more likely to be promoted. For the top panel, we compute residualized collaborators, equal to the residual from a regression of the log number of collaborators on log sales and all other control variables in Column 4 of Table 3. We then plot the distribution of promotion propensity separately for workers with residualized collaboration experience above or below the median. The overlapping distributions suggest that firms are only more likely to promote workers with more collaboration experience insofar as that is positively correlated with other variables. Controlling for these other factors, employees with more collaboration experience are no more likely to be promoted.

3 What predicts managerial performance?

Next, we examine the relation between pre-promotion worker characteristics and post-promotion manager value added:

$$\begin{aligned} \text{Manager Value Added}_{if} &= b_1 \text{Pre-Promotion Sales}_{if} + b_2 \text{Pre-Promotion Collaborators}_{if} \\ &\quad + W_{if} + u_{if} \end{aligned} \tag{3}$$

We estimate Equation (3) at the manager level because manager value added is defined as a time-invariant manager characteristic. Pre-Promotion Sales$_{if}$ is the log of one plus manager i's monthly sales credits as a worker, averaged over the 12 months prior to i's promotion (or for i's total tenure pre-promotion tenure, if less than 12 months). Analogously, Collaborators$_{if}$ is the log of manager i's number of distinct collaborators each month, averaged over the year prior to promotion (or for i's total pre-promotion tenure, if less than 12 months). In some specifications, we also control for manager i's tenure and team size in the month prior to the promotion event.

[Table 5]

Table 5 shows that, among promoted managers, there is a significant *negative* relation between

pre-promotion sales performance and subsequent managerial performance. Column 2 shows that a doubling in a manager's pre-promotion sales corresponds to an approximate 10 percentage point decline in manager value added. Since manager value added represents the change in log subordinate sales, this implies that a manager with double the pre-promotion sales leads each subordinate's sales to decline by 10 percent. Given that a typical manager is in charge of five subordinates, our results also imply that a doubling of a manager's pre-promotion sales predicts that total team sales under the new manager will decline by approximately half of one worker. In contrast, collaboration experience is positively correlated with manager value added. In Column 2, we find that a doubling in collaboration experience predicts a large 24 percentage point improvement in manager value added.

In some cases, workers are promoted to replace their former managers. In these types of promotions, one may be concerned that our estimates of the relation between manager value added and pre-promotion characteristics may be driven by direct changes in work assignments within the team. The promotion of stronger salespeople on a team may adversely impact remaining subordinates by removing valuable accounts that they may have also been contributing too. In this case, we may see a mechanical correlation between pre-promotion sales and the subsequent performance of remaining subordinates.

To ensure that such possibilities are not driving our results, we further restrict our sample to promoted managers that are assigned to subordinates that were not also their former team members. Specifically, we require that more than two-thirds of the newly promoted manager's assigned subordinates were not previously assigned to the same parent (manager) as the newly promoted manager prior to her promotion. In Columns 3 and 4 of Table 5, we find similar results in this restricted sample.

3.1 Correcting for selection

The empirical results so far show that firms promote based on current job performance even though pre-promotion sales negatively predicts managerial performance and other observable characteristics positively predict managerial performance. This evidence is consistent with the Peter Principle

and the idea that firms fail promote the best potential managers. However, we face the empirical challenge that we do not observe latent managerial potential for all workers. Instead, we observe managerial performance for the subset of workers that are actually promoted. Although we estimate a negative relation between sales and managerial performance for the subset of promoted workers, it is possible that the overall average relation between sales and latent managerial performance is positive. If this were the case, then emphasizing sales in promotion decisions could be an effective way to promote workers with the greatest managerial potential. To evaluate the Peter Principle, we need to estimate the relation between sales and managerial performance for the full sample of workers, both promoted and not promoted.

To address this measurement challenge, we apply a two-stage selection model in the style of Heckman (1976), Heckman (1979), Grabner (1974), and Lewis (1974). The goal of this selection correction is to recover the predictive relation between sales performance and latent managerial potential for the full sample of workers, so that we can assess whether it is really the case that firms promote high performing sales workers even though sales performance negatively predicts managerial performance.

Suppose that the underlying relation between latent managerial potential M_i and worker characteristics is given by:

$$M_i = \beta_1 \text{Sales}_i + \beta_2 \text{Collaborators}_i + X_i\beta_3 + \varepsilon_i. \tag{4}$$

However, we only observe M_i if workers are promoted. Firms follow a promotion rule given by:

$$P_i = \mathbb{I}\left(\tau_1 \text{Sales}_i + \tau_2 \text{Collaborators}_i + X_i\tau_3 + Z_i\tau_4 + \mu_i > 0\right) \tag{5}$$

This promotion rule is flexible and allows promotion to depend on observable worker characteristics that impact managerial potential ($Sales_i$, $Collaborators_i$, and X_i) as well as other factors Z_i that affect the probability of promotion but not managerial potential. We do not impose any restrictions on the relation between Equations (4) and (5) other than the standard assumption in Heckman-style selection models that their respective errors terms are jointly normally distributed with mean zero and correlation ρ, and that there are exist variables Z_i that impact promotion but do not relate to

managerial performance (discussed shortly).

Our empirical argument for the Peter Principle centers on showing that both $\tau_1 > 0$ (firms are more likely to promote better sales performers) and $\beta_1 < 0$ (better sales performers are more likely to be worse managers). If this were the case in the full sample of workers, it must be that better sales performers are receiving some kind of positive boost in promotion probability unrelated to their managerial potential. We have already shown $\tau_1 > 0$ in the full sample of workers, and we now seek to recover β_1 after correcting for sample selection.

Before continuing, we note that $\tau_1 > 0$ and $\beta_1 < 0$ is a sufficient, but not necessary criterion for the Peter Principle. Firms may be promoting the best sales workers at the expense of managerial match quality even if the underlying relation between sales and latent managerial performance were positive. For example, suppose that both sales and collaboration experience were positively related to latent managerial performance. Firms may still weight sales performance too much in promotion decisions relative to collaboration experience, and thereby overemphasize current job performance in promotion decisions at the expense of promoting the best potential managers. In such cases, promotions decisions "overweight" sales performance relative to its true relationship with managerial potential. In other words, the Peter Principle could still apply if $\tau_1 > 0$ and $\beta_1 > 0$, but τ_1 were too large relative to the weights on other characteristics in the promotion rule. In practice, we will show that sales performance continues to negatively predict managerial performance after correcting for selection, so any positive weight on sales in the promotion decision represents a promotion policy that does not maximize managerial match quality.

In our setting, latent managerial performance defined in Equation (4) is observed only when workers are promoted according to the rule defined in Equation (5). Following the standard Heckman correction procedure, we first estimate worker promotion propensity in the full worker-month level panel using a probit model:

$$\begin{aligned}\Pr(P_{ift} = 1 | \text{Sales}_{ift}, \text{Collaborators}_{ift}, X_{ift}, Z_{ift}) \\ = \Phi(a_1 \text{Sales}_{ift} + a_2 \text{Collaborators}_{ift} + X_{ift} a_3 + Z_{ift} a_4) \quad (6)\end{aligned}$$

This first stage is identical to the probit regression model that was already presented in Columns 3 and 4 of Table 4. We then recover estimates for the β's in Equation 4 by estimating a regression similar to Equation 3 (presented earlier), also controlling for the inverse Mills ratio, m_i, of the fitted values from the first stage regression:

$$M_i = b_1 \text{Sales}_i + b_2 \text{Collaborators}_i + X_i b_3 + b_4 m_i + e_i. \quad (7)$$

This selection model allows us to recover unbiased estimates for the βs in Equation (4). Here, the inverse Mills ratio is a function of sales, collaborations, and other covariates X and Z. Crucially, we assume that there exist variables Z_i that impact a worker's probability of promotion, but not her managerial potential. If this assumption holds, then the impact of m_i is separately identified from the other variables in the second stage regression because we assume that Z_i does not directly impact managerial performance and thus would not directly enter into this regression. This is equivalent to saying that the Z_i's are instruments for promotion.[7]

For Z_i, we use industry-level and firm-level average promotion rates or firm-year-month fixed effects, which subsumes the firm- and industry-level promotion rates within each period. One may be concerned that differences in industry or company conditions captured by promotion rates or company-time fixed effects may also reflect strong consumer demand and other firm-level factors, which may directly impact the performance of managers. However, we measure managerial quality as the value aded of managers to subordinate sales, after controlling for worker and firm-year-month fixed effects. As such, for our observed sample of managers, our measure of manger quality is, by construction, orthogonal to our instruments for selection, which only vary at the firm by year-month level.

[Table 6]

Formally, our instruments must be orthogonal to latent managerial potential in the full set of

[7]The selection correction and the use of the Z_i instruments allows us to recover an underlying predictive relation in the full sample of workers, not to establish causality. It remains possible that something correlated with sales performance, rather than sales performance per se, causes lower managerial performance. Nevertheless, we argue that firms, by allowing sales performance to positively predict promotion when it negatively predicts managerial performance, do not promote the best potential managers.

workers. While it is impossible to fully test this exclusion restriction because we do not observe latent managerial performance for the full sample of workers, we can test whether industry and firm-level promotion rates in each year month are correlated with manager value added for the sample of promoted managers. In Table 6, we show that for both the sample of all promotions, and promotions to different teams, we find that the instruments are uncorrelated with manager value added. This is unsurprising given that we measure manager value added as the manager fixed effects extracted from a regression that also controls for firm-year-month fixed effects. In fact, one would expect the point estimates on the relationship between manager value added and our instruments to be precisely zero; the reason they can potentially differ here is because we estimate manager value added on a larger set of managers. Specifically, to estimate value added, we include both managers whom we only ever observe as managers, as well as those who were once workers and whom we see promoted; for the former set, we do not observe their pre-promotion sales performance. Table 6 simply verifies that there is no correlation between observed manager value added and our instruments, for our analytic set of managers for whom we observe pre-promotion performance.

[Table 7]

Table 7 reports the relation between pre-promotion characteristics and manager value added, corrected for selection. We continue to find that sales performance negatively predicts manager value added. Using firm and industry-level promotion rates as the selection instruments in Column 1, we estimate that a 10 percent increase in pre-promotion sales performance predicts a 1 percent decline in the sales performance of subordinates. Using firm-year-month fixed effects as the selection instruments in Column 2, the effect is larger: a 10 percent increase in pre-promotion sales performance predicts a 2 percent decline in the performance of subordinates. Overall, these estimates are similar in magnitude and slightly larger in absolute magnitude that those in Table 5 which did not correct for selection issues.

We also continue to find that collaboration experience positively predicts manager value added. The estimated magnitudes are economically large, but significant only when using promotion-rates rather than firm-year-month fixed effects. Columns 1 and 2 show that a 10 percent increase in

pre-promotion collaboration experience predicts a 2.4 and 1.6 percent increase in subordinate performance, respectively. In Columns 3 and 4, we again restrict the sample to promoted managers that are assigned to subordinates that were not also their former team members. We find similar results in this restricted sample, suggesting that our findings are not driven by unusual time trends occurring when workers are promoted to replace their former managers.

[Figure 6]

These results suggest that firms "discriminate" in favor of strong salespeople by lowering their promotion thresholds or holding biased beliefs about the relation between sales and managerial performance. An implication of this behavior is that, among promoted managers with the same propensity to be promoted, weaker salespeople should be more successful. Figure 6 provides direct evidence for this prediction. The top panel plots manager value added versus the estimated promotion propensity of promoted managers. The blue line represents managers from the bottom half of pre-promotion sales distribution while the red line plots the same relation for managers from the top half of the pre-promotion sales distribution, controlling for company fixed effects. The vertical distance between these lines is consistent with discrimination: given the same propensity to be promoted, managers with lower pre-promotion sales performance consistently have higher managerial performance. This difference is not statistically significant for lower promotion propensities, but among managers who were very likely to be promoted, we see a large difference. This is consistent with a story in which firms are very likely to promote the best sales workers, even though these workers are unlikely to be good managers. If we see poor sales workers who are very likely to be promoted, it must be the case that they had exceptionally high managerial potential.

4 Potential alternative explanations

Lazear (2004) shows that mean reversion can generate patterns that, on the surface, look like the Peter Principle. The highest performing sales workers at any given point in time may be ones who are currently selling above their individual mean; if they are promoted at this point, their performance as managers may fall as a result of mean reversion. Mean reversion implies that even

if firms were promoting the best potential managers, we may still see a decline in within-person performance after promotion. Our results cannot be explained by this type of mean reversion for two fundamental reasons. First, our measure of managerial performance is not based on a manager's own sales, but is instead based on the value-added to the sales performance of his or her subordinates. Thus, mean reversion in the manager's own sales performance does not affect our measure of manager value added. Second, mean reversion cannot explain the negative cross-sectional relation between pre-promotion sales and managerial performance. If better sales workers do indeed make better managers, mean reversion should only attenuate this positive correlation. It would not predict a negative cross-sectional relation between pre-promotion sales and post-promotion managerial performance, as we find in the data. In our sample, firms could promote better managers by simply selecting workers with lower sales performance, all else equal.

Another potential concern is that newly promoted managers may be assigned to subordinates in a non-random manner. In general, a simple correlation between the pre-promotion sales of newly-promoted managers and the *level* of performance of their assigned subordinates should not impact our results because we estimate manager value added from changes in subordinate performance under the new manager. However, one may still be concerned about manager pre-promotion performance being correlated with time-varying aspects of a worker's sales performance. In particular, managers with high pre-promotion sales may be systematically assigned to subordinates whose sales are likely to decrease after being assigned to that manager, for reasons unrelated to that manager.

[Table 8]

Table 8 explores the nature of manager-subordinate assignment. The first key pattern that emerges is that managers are not randomly assigned to subordinates: a 10 percent increase in a manager's pre-promotion sales is correlated with an approximate 4 percent increase in the prior sales of the subordinates to whom he or she is assigned. However, managers with higher pre-promotion sales do not appear to be assigned to subordinates with different time trends in performance. Table 8, Columns 1, 2, and 3 examine subordinates' sales in the 7-12 months, 4-6 months, and 1-3 months prior to the arrival of the new managers, and finds no evidence of either a positive or negative time trend associated with the new manager's pre-promotion sales performance. The stability of these

estimates suggests that managers with higher pre-promotion sales are not assigned to subordinates with increasing or decreasing trends in performance. We also find that manager's pre-promotion collaboration experience does not significantly predict assignment to subordinates.

We also explore potential biases from non-random assignment by considering how subordinates' sales prior to a manager's promotion are correlated with the new manager's estimated value added. For example, high performing sales workers may have less scope for improvement. If so, managers who are assigned to high prior sales subordinates may have low value added simply because these subordinates are already such high performers. In this case, manager value added should be negatively correlated with a subordinates' prior sales. Table 8, Columns 4, 5, and 6 instead shows a statistically insignificant and inconsistently signed relation between subordinate prior sales and manager value added.

One may also be concerned about a different type of selection issue in which some top sales workers prefer not to be promoted. Although most workers enjoy significant pay increases after promotion, the very top sales workers in our sample earn more than the typical sales manager. It may be the case that some top sales workers do not want to be promoted, and as a consequence, we do not observe managers with very high pre-promotion sales in our sample of promoted workers. This is likely to be a bias against our findings that higher pre-promotion sales is associated with lower manager value added. Sales workers who are offered promotions should compare their expected pay as managers with their expected pay as sales workers, and then decide whether to accept the promotion. Thus, workers with strong sales should only accept promotions if they have very good prospects as managers. In other words, the selection in terms of who accepts promotion should bias toward finding that better sales workers make better managers, contrary to our finding that better sales workers become worse managers.

Even if it were the case that the very best sales workers actually make good managers, but prefer to remain in their current roles, our results still indicate that firms are not maximizing managerial performance by promoting good sales workers. If firms wish to maximize managerial match quality, they should promote the best potential managers from the set of workers who would be willing to accept such a promotion.

[Table 9]

In Table 9, we explore potential non-linearities in the relation between sales performance and promotion probability. We also explore non-linearities in the relation between sales performance and manager value added. An inverted relation for certain parts of the sales distribution would suggest that promoting based on sales may actually help to maximize managerial match quality, at least within certain parts of the sales distribution. For example, suppose that pre-promotion sales performance negatively predicts manager value added *on average*, but positively predicts manager value added for high values of sales. If so, firms may be maximizing managerial match quality when they promote based on sales performance, at least among the sample of workers with high sales.

We regress a dummy variable for whether a worker is promoted on a three part spline (Column 1) and five part spline (Column 2) for sales performance. The knots in the spline are spaced to equally divide the sample into three or five groups, and the coefficients represent the slope of the relation between promotion probability and sales within each sales interval. Similarly, Columns 3 and 4 show the second stage regression in a Heckman model of manager value added on a three and five part spline with respect to pre-promotion sales. The coefficients again represent the slope of the relation between manager value added and sales within each sales interval.

Empirically, we do not find strong evidence of inversions. The estimated magnitudes sometimes become insignificant once we examine subsamples of the sales performance distribution. However, we consistently find a negative relation between pre-promotion sales and manager value added across all intervals in the sales distribution. We also find that promotion probability is low and insensitive to sales in intervals where sales is low, but then strongly positively related to sales for sales in tercile 2 and above or quintile 3 and above.

These patterns are underscored by Figures **??** and **??**, which plot the relationship between pre-promotion sales and collaborations on promotion propensities and manager value added, respectively. In each case, we observe a monotonic relationship.

5 What are the performance losses from mismatch?

How much do the observed promotion policies cost firms in terms of lost subordinate sales? To analyze this, we set aside tournament incentives, monitoring constraints, fairness concerns, and other potential benefits of firms' promotion rules to focus instead on the costs of managerial mismatch. If firms do not make mistakes in promotion policies, then our estimates may be interpreted as the match quality that firms forgo to achieve better incentives.

Specifically, we examine how predicted managerial performance differs among three categories of workers: (1) actual promoted salespeople, (2) non-promoted salespeople among the promoted worker's peers, and (3) the top predicted manager among a promoted sales worker's peers. Peers are defined as other salespeople in a team managed by the same manager in the same time period. We interpret case (3) to be the performance-maximizing promotion decision under the restriction that mobility and other frictions prevent the firm from promoting among the entire organization, and rather, firms must promote among the peers of promoted workers. If we relax this restriction, then the estimated costs of mismatch will further increase. If firms face greater restrictions in promotions than we impose in our simulation, then we would overestimate the costs of mismatch.

[Figure 7]

We predict manager value added for each group using fitted values from the regression presented in Table 7 Column 2. Figure 7 shows the distributions of predicted manager value added in these three samples. The mean predicted improvement in subordinate salesperson performance is scaled to zero for the sample of promoted workers. The mean change in subordinate sales performance among the sample of non-promoted workers is -0.13, implying that firms current promotion policies do better than promoting at random. The mean in the sample of best predicted managers is 0.29, implying that subordinate performance could improve by up to 30 percent if firms pursued an alternative promotion policy of promoting the best predicted manager within a sales team.[8]

[8] As discussed previously, the variance of manager value added may be biased upwards because of measurement error in the estimation of the manager fixed effects. However, the measurement error associated with manager value added should not bias regression coefficients and estimates in this counterfactual, which uses fitted values from the regression presented in Table 7 Column 2.

6 Variation by in pay-for-performance

We now explore how promotion policies vary with pay-for-performance. Baker, Jensen and Murphy (1988) suggest that pay-for-performance incentivizes worker effort and may offer a substitute for the tournament incentives associated with promotion policies that prioritize current job performance.

We estimate firm-level pay-for-performance as the ratio of commissions plus bonus compensation scaled by base salary, averaged across all workers in the firm within each calendar year. Before proceeding, we note some advantages and disadvantages of our data with regard to compensation. An advantage of our data is that we observe commission and bonus compensation for a subset of workers across a large number of firms, allowing us to study heterogeneity across firms. However, we only observe compensation data for a subsample of firms, leading to an approximate 35 percent decline in sample size. Further, as noted previously, our data provider's software is not designed to track or distribute base salaries, and instead treats base salaries as an optional field. Therefore, salary data can be missing or measured with error (for example, if it's not updated or the pay periods are unclear). Nevertheless, we believe our estimates of fraction variable pay to be a reasonable, if noisy, proxy for the strength of pay-for-performance incentives across firms in our data.

[Table 10]

In Table 10, we regress whether a worker is promoted on the interaction between worker sales and collaboration experience and the fraction variable pay associated with each firm year. We also control for all direct effects. We find that firms that grant more performance-sensitive pay relative to fixed pay also implement promotion policies that are less sensitive to worker sales performance. This is consistent with the idea that pay-for-performance incentives can partially offset the need to provide incentives through promotion tournaments based on sales performance. However, our results do not necessarily imply that pay-for-performance can eliminate the costs associated with the Peter Principle. Pay-for-performance may be an expensive substitute for tournament incentives from the point of view of firm management and shareholders. Further, relative to other occupations, sales is associated with very high pay-for-performance on average, and yet we continue to find evidence of the Peter Principle within the sales occupation.

7 Conclusion

We use detailed microdata on the performance and promotions of sales workers at a large number of firms to test the adage that "the best salesperson doesn't always make the best manager." We find strong empirical support for the Peter Principle. The most productive worker is not the best candidate for manager, and yet firms are significantly more likely to promote top front line sales workers into managerial positions. In making promotion decisions, firms weigh sales performance more and collaboration experience less than they should if they were trying to maximize managerial match quality. As a result, the performance of a new manager's subordinates decline relatively more after the managerial position is filled by someone who is a strong salesperson, but typically worked alone, prior to promotion.

We caution against interpreting these results as evidence that firms have mistaken beliefs or behave inefficiently. Rather, consistent with tournament theory, firms may heavily weigh sales performance in promotion decisions to encourage workers to exert effort in their current job roles and to maintain norms of fairness. In addition, the availability of relatively clear measures of worker productivity among frontline sales workers may lead organizations to emphasize these characteristics rather than other, more subjective or fungible employee characteristics in promotion decisions. Regardless, our results lend evidence that firms do not promote entirely to maximize match quality. Our results also imply that managerial match quality, tournament incentives, and other objectives are not perfectly aligned.

References

Abowd, John M, Francis Kramarz, David N Margolis, and Kenneth R Troske. 2001. "The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States." *Journal of the Japanese and International Economies*, 15(4): 419–436.

Baker, George P, Michael C Jensen, and Kevin J Murphy. 1988. "Compensation and Incentives: Practice vs. Theory." *The Journal of Finance*, 43(3): 593–616.

Deutsch, Claudia. 1986. "Holding on to Technical Talent." *New York Times*.

DeVaro, Jed, and Oliver Gürtler. 2015. "Strategic Shirking in Promotion Tournaments." *Journal of Law, Economics, and Organization*, ewv019.

Fairburn, James A, and James M Malcomson. 2001. "Performance, Promotion, and the Peter Principle." *The Review of Economic Studies*, 68(1): 45–66.

Grabner, Isabella, and Frank Moers. 2013. "Managers' Choices of Performance Measures in Promotion Decisions: An Analysis of Alternative Job Assignments." *Journal of Accounting Research*, 51(5): 1187–1220.

Grabner, R. 1974. "Wage Comparisons: A Selectivity Bias." *Journal of Political Economy*, 82: 1119–1143.

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement*, 51: 475–492.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica: Journal of the Econometric Society*, 153–161.

Kaplan, Steven N, and Morten Sorensen. 2016. "Are CEOs Different? Characteristics of Top Managers." *Working paper*.

Kaplan, Steven N, Mark M Klebanov, and Morten Sorensen. 2012. "Which CEO Characteristics and Abilities Matter?" *The Journal of Finance*, 67(3): 973–1007.

Lazear, Edward P. 2004. "The Peter Principle: A Theory of Decline." *The Journal of Political Economy*, 112(S1): S141–S163.

Lazear, Edward P, and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *The Journal of Political Economy*, 89(5): 841–864.

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton. 2016. "Who Gets Hired? The Importance of Finding an Open Slot." National Bureau of Economic Research.

Lewis, H. G. 1974. "Comments on Selectivity Biases in Wage Comparisons." *The Journal of Political Economy*, 21: 1145–1155.

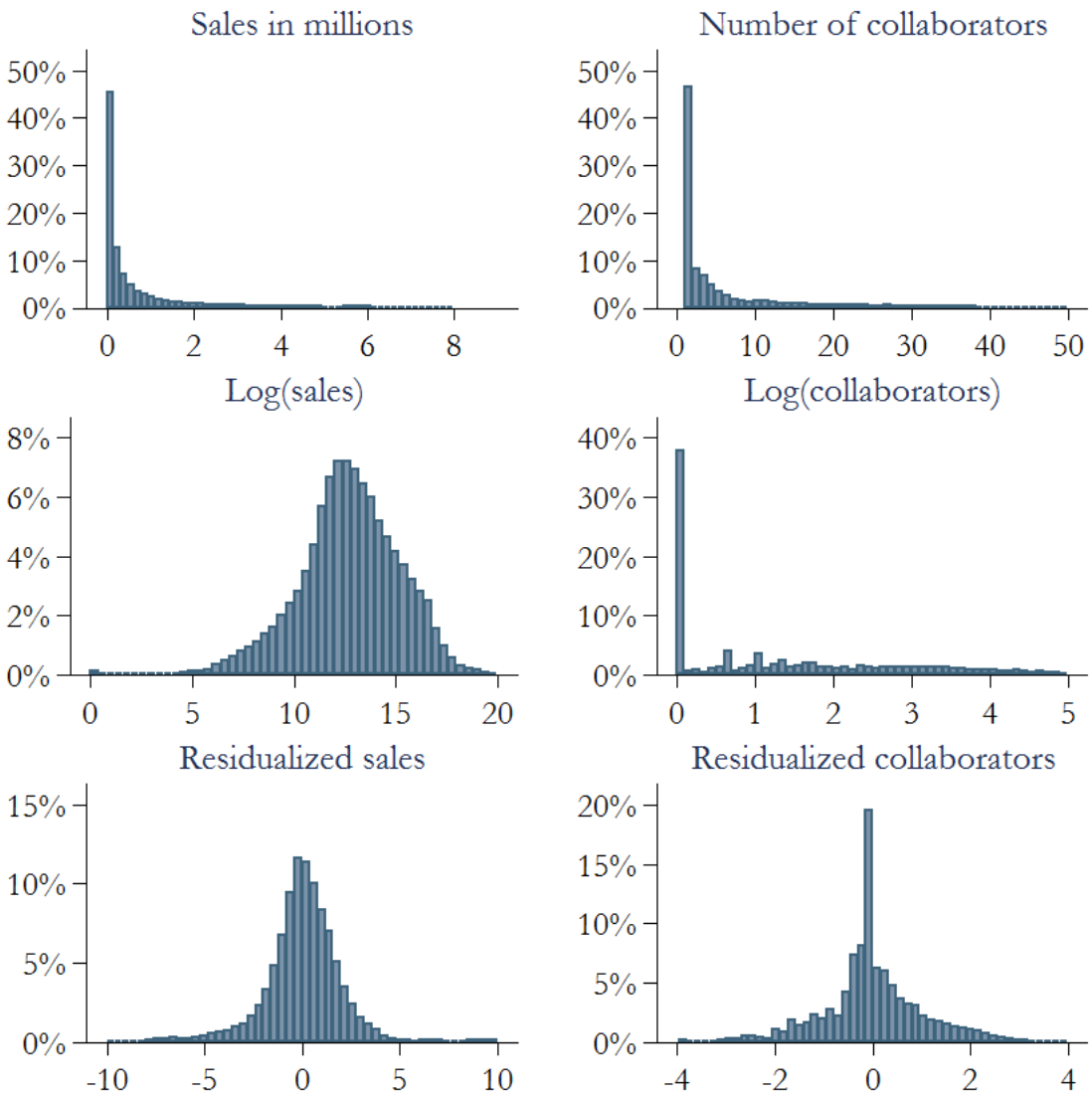
Murphy, Kevin J, and Jan Zabojnik. 2004. "CEO Pay and Appointments: A Market-Based Explanation for Recent Trends." *The American Economic Review*, 94(2): 192–196.

Peter, Laurence J, and Raymond Hull. 1969. *The Peter Principle.* New York: William Morrow & Co.

Prendergast, Canice. 1998. "What Happens within Firms? A Survey of Empirical Evidence on Compensation Policies." In *Labor Statistics Measurement Issues*. 329–356. University of Chicago Press.

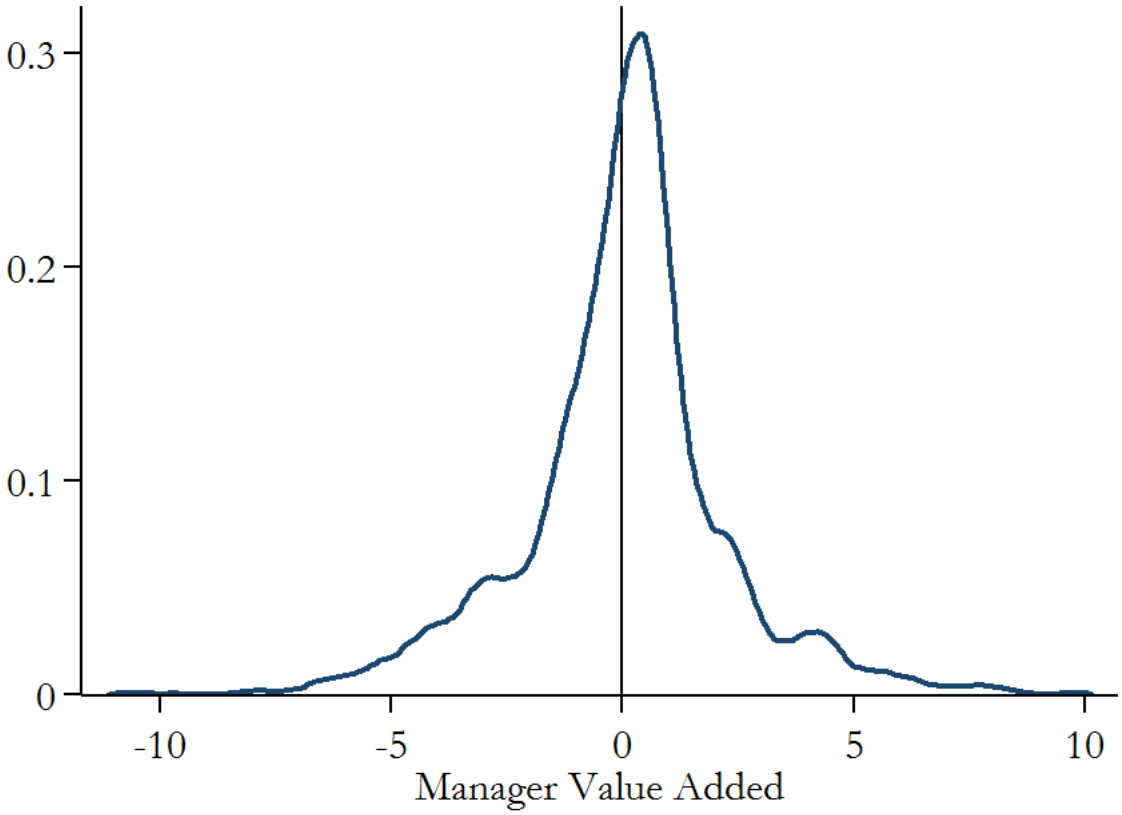
Waldman, Michael. 2003. "Ex Ante Versus Ex Post Optimal Promotion Rules: The Case of Internal Promotion." *Economic Inquiry*, 41(1): 27–41.

Figure 1: Distribution of sales and number of collaborators



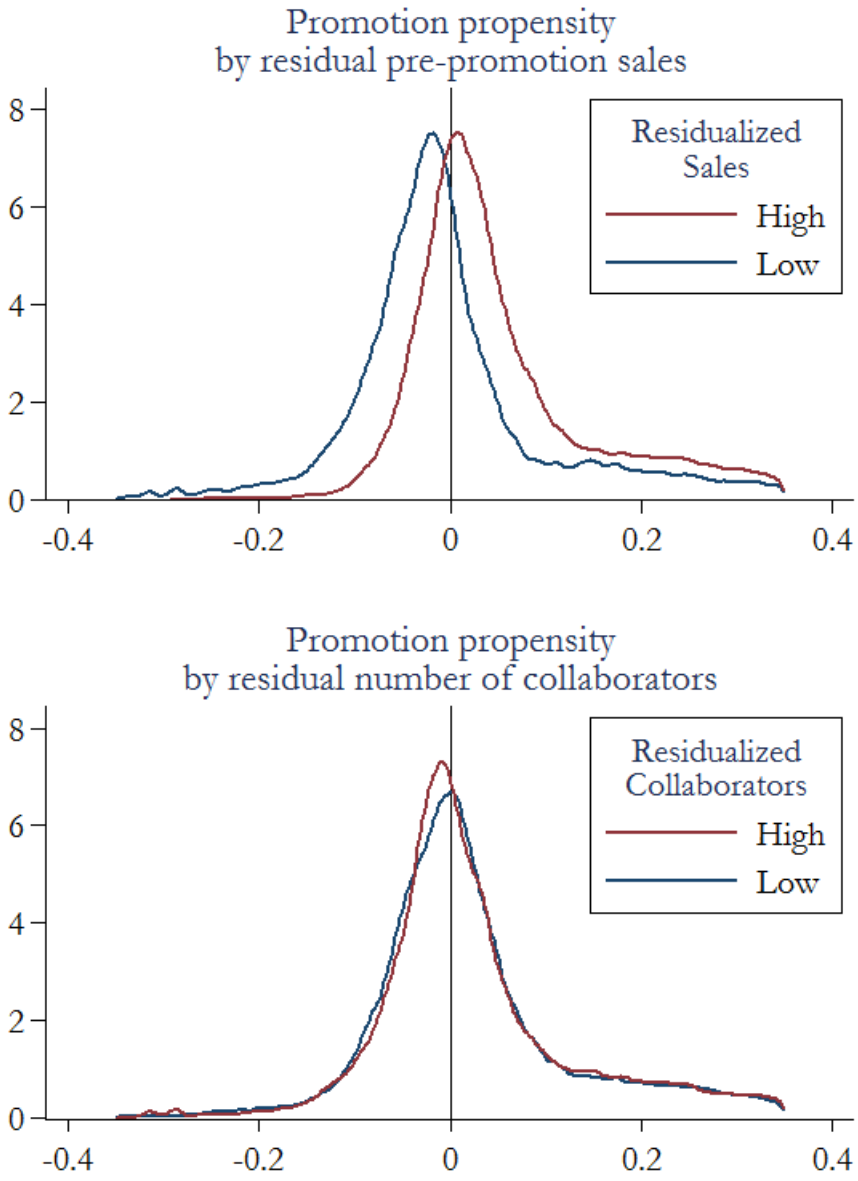
Notes: Left and right panels are 12-month moving averages for sales and number of collaborators, respectively. The top panels show the untransformed distribution. The middle panels show the log-transformed distribution. The bottom panels show the residuals after the log-transformed variables are regressed on firm-year-month fixed effects.

Figure 2: Distribution of Manager Value Added



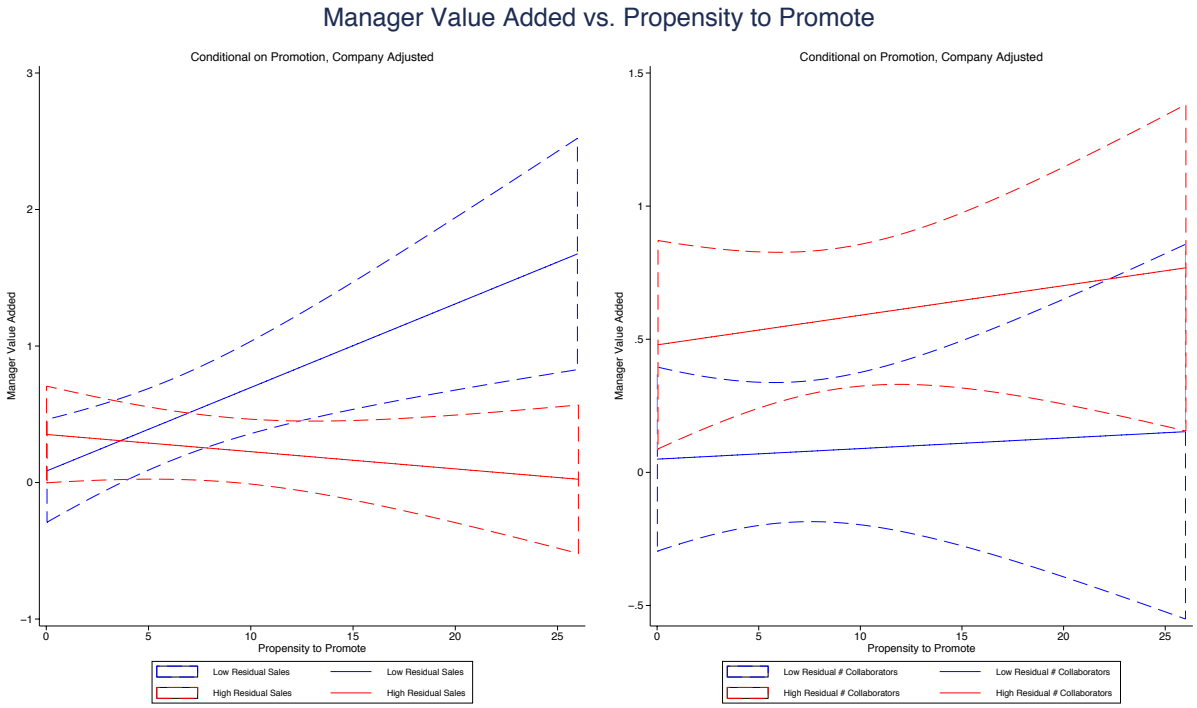
Notes: The figure shows the kernel density of estimated manager value added.

Figure 3: Distribution of Promotion Propensity



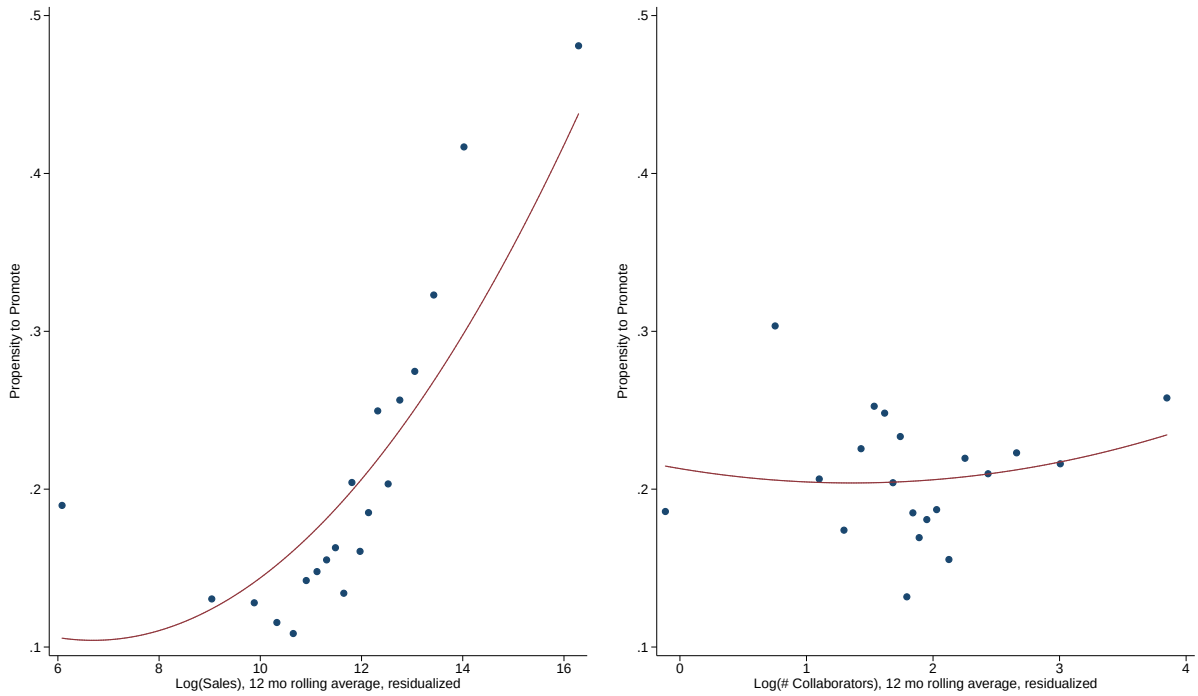
Notes: The top panel shows the kernel density of the predicted values of promotion propensity for salespeople above and below the median for residualized sales, where residualized sales is estimated as the residual from a regression of log sales on firm-year-month fixed effects. The bottom panel shows the kernel density of the predicted values of promotion propensity for salespeople above and below the median for residualized collaborators, where residualized collaborators is estimated as the residual from a regression of log number of collaborators on firm-year-month fixed effects.

Figure 4: Distribution of Manager Value Added



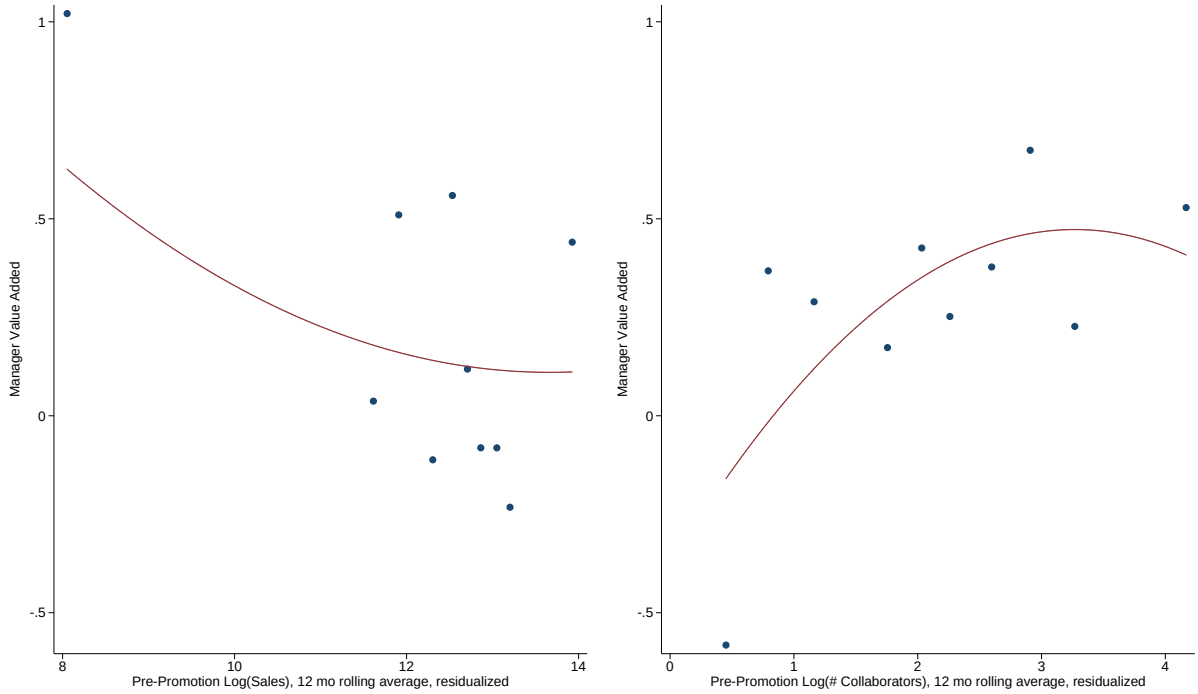
Notes: The figure shows the relation between estimated manager value added among observed promotions (the y-axis) and the estimated promotion propensity (the x-axis). The left panel shows this relation separately for the sample with above/below median residualized sales (after controlling for firm fixed effects). The right panel shows this relation separately for the sample with above/below median residualized number of collaborators (after controlling for firm fixed effects). The dashed lines represent 90% confidence intervals.

Figure 5: Pre-promotion characteristics and promotion propensity



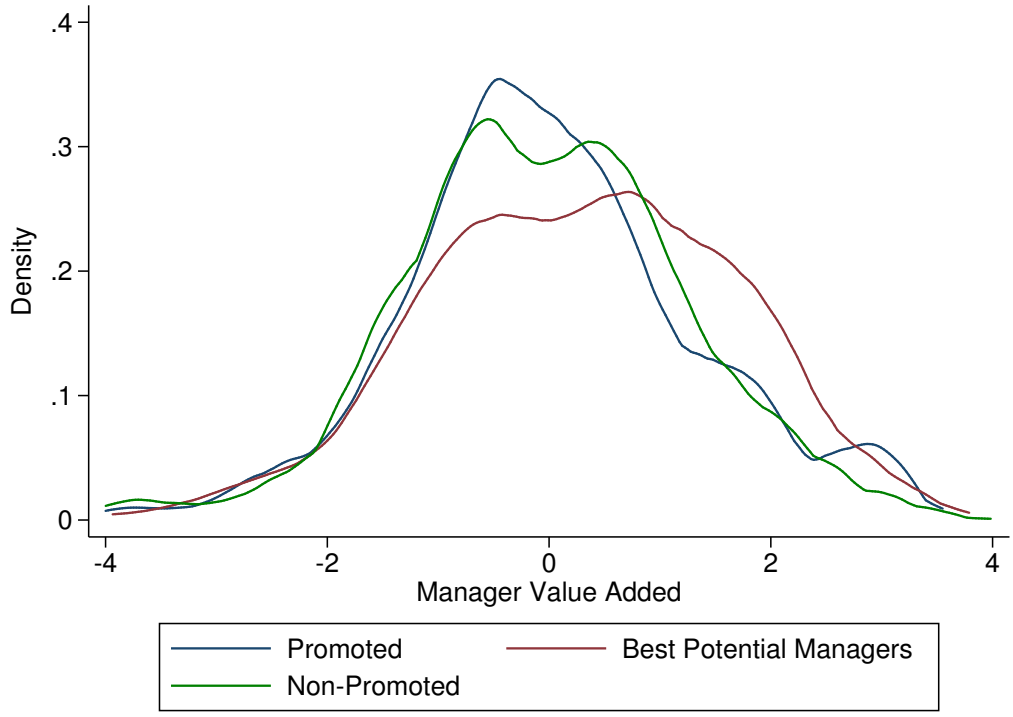
Notes: The left panel plot in this figure shows a bin scatter relation between pre-promotion sales (x-axis) and estimate promotion propensities among observed promotions (the y-axis). The left panel is residualized to control for pre-promotion collaborations, tenure and team size dummies. The right panel plots this same relationship for pre-promotion collaborations. In this case, pre-promotion collaborations is residualized to control for pre-promotion sales, tenure and team size dummies. Each panel plots a quadratic line of best fit to allow for the possibility of inversions in the estimated relationship.

Figure 6: Pre-promotion characteristics and manager value added



Notes: The left panel plot in this figure shows a bin scatter relation between pre-promotion sales (x-axis) and estimated manager value added among observed promotions (the y-axis). The left panel is residualized to control for pre-promotion collaborations, tenure and team size dummies as well as the hazard of selection into the promoted sample. The right panel plots this same relationship for pre-promotion collaborations. In this case, pre-promotion collaborations is residualized to control for pre-promotion sales, tenure and team size dummies, and the hazard of promotion. Each panel plots a quadratic line of best fit to allow for the possibility of inversions in the estimated relationship.

Figure 7: Actual vs. Optimal Promotion Policies



Notes: The figure shows the distribution of predicted manager value added for the samples of workers that are promoted, workers that are not promoted, and workers that would make the best potential managers. All manager value added measures are predicted fitted values of manager value added based on the results in Table 7 Column 2, using covariates at the time of the promotion event. To determine the best potential managers sample, we select the worker with the highest predicted manager value added within the same team and month when a worker is actually promoted. The non-promoted sample consists of other, non-promoted, workers in the same team and month when a worker is actually promoted.

Table 1: Descriptive statistics

Sample coverage		**Pr(Promotion)**	
# Firms	214	Overall	0.0298
# Workers	48,209	Monthly hazard	0.0021
# Managers	5,369		
# Promotions	1,565		
Years covered	2005-2011		

Summary statistics	*Mean*	*25th*	*50th*	*75th*
Worker characteristics				
Monthly sales*	$3,573,065	$42,883	$272,987	$1,498,726
# Collaborators*	21.6	1	3	14
Monthly commissions*	$12,358	$1,048	$3,584	$ 9,598
Salary	$7,007	$4,294	$6,911	$9,167
Manager characteristics				
# of subordinates	4.6	2	4	6
Monthly commissions*	$15,014	$2,509	$6911	$16,524
Change in monthly commissions	$1,529	-$1,157	$908	$5,834
Salary	$12,439	$8,855	$11,538	$14,062

Notes: * denotes 12 month moving average. The change in monthly commissions represents changes in pay after promotion, estimated as average monthly commissions in the 12 months after promotion minus average monthly commissions in the 12 months before promotion.

Table 2: Manager samples and firm-level means

Manager sample size	
# of managers	5,369
# of managers with mover subordinates	2,569
# of managers with estimated fixed effects	2,231
# of internally promoted managers with estimated fixed effects	696

Firm-level means	*Mean*	*25th*	*50th*	*75th*
Share of workers who switch managers	45.0	20.9	44.4	72.7
Average size of connected group (worker-months)	16,313	2,862	7,104	23,509
Share of workers in largest connected group	88.5	85.3	96.5	99.3

Notes: Managers with mover subordinates are managers with at least one subordinate who has worked under other managers within our data sample. Internally promoted managers are managers for whom we observe as workers prior to promotion.

Table 3: Linear Probability Model for Promotions

	Worker is promoted			
	(1)	(2)	(3)	(4)
Log(sales)	0.0181***	0.0225***	0.0177***	0.0329***
	(0.00180)	(0.00184)	(0.00178)	(0.00292)
Log(collaborators)	0.0157***	0.0231***	0.00481	0.00934
	(0.00435)	(0.00436)	(0.00428)	(0.00719)
Industry-month promotion rate		1.001***	0.00524	
		(0.0817)	(0.0561)	
Firm-month promotion rate			0.998***	
			(0.0409)	
Pre-promotion characteristics	Yes	Yes	Yes	Yes
Firm-month FE	No	No	No	Yes
R-squared	.001	.01	.056	.056
Observations	734278	734278	734278	734278

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the regression described in Equation (2). We use data at the worker-month level for workers that have not yet been promoted. The dependent variable is an indicator for whether a worker is promoted in the next month. Log sales is the log of one plus worker i's monthly sales credits, averaged over the past 12 months or for the worker's total tenure if tenure is less than 12 months. Log collaborators is the log of one plus the unique number of sales collaborators in a month who have shared credits with worker i, again averaged over the past 12 months or for the worker's total tenure if tenure is less than 12 months. Pre-promotion characteristics include controls for seven bins of a worker's tenure (and tenure interacted with a censored dummy) and seven bins for team size. Team size is distinct from collaboration experience and represents the number of other sales workers who share the same manager (team members do not necessarily collaborate on sales transactions). Industry-month promotion rate is the percentage of workers promoted within the an industry-year-month. Firm-month promotion rate is the percentage of workers promoted within worker i's firm in the same month. Column 4 includes controls for firm-year-month fixed effects, which subsumes the controls for industry-year-month and firm-year-month promotion rates. Coefficients in this table only have been multiplied by 100 to improve readability and represent the percentage point change in promotion probability for a unit change in each independent variable. Standard errors are allowed to be clustered by worker.

Table 4: Probit Model for Promotions

	Worker is promoted			
	(1)	(2)	(3)	(4)
Log(sales)	0.0314***	0.0336***	0.0309***	0.0574***
	(0.00348)	(0.00357)	(0.00399)	(0.0112)
Log(collaborators)	0.0232***	0.0287***	0.0206***	0.0243
	(0.00577)	(0.00581)	(0.00660)	(0.0169)
Industry-month promotion rate		0.135***	-0.0535**	
		(0.00596)	(0.0233)	
Firm-month promotion rate			0.125***	
			(0.00308)	
Pre-promotion characteristics	Yes	Yes	Yes	Yes
Firm-month FE	No	No	No	Yes
Observations	734142	734142	734142	734142

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table estimates the same specifications as in Table 3 using a probit model. Reported coefficients are those from the probit model and do not represent marginal effects. All variables and sample restrictions are as defined for Table 3. Standard errors are allowed to be clustered by worker.

Table 5: Predictors of Manager Value Added

Manager value added	All positions		Promoted to different team	
	(1)	(2)	(3)	(4)
Log(sales)	-0.0864**	-0.103**	-0.0977***	-0.122**
	(0.0354)	(0.0495)	(0.0364)	(0.0499)
Log(collaborators)	0.258**	0.239**	0.252**	0.222*
	(0.111)	(0.111)	(0.120)	(0.120)
Pre-promotion controls	No	Yes	No	Yes
R-squared	.033	.068	.041	.101
Observations	696	696	608	608

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the regression described in Equation (3). We use data at the manager level. The sample is restricted to promoted managers for whom we can observe pre-promotion characteristics and for whom we can estimate manager value added fixed effects using movements of subordinates across managers. The dependent variable is manager value added, estimated as the change in subordinate performance associated with each manager (see Equation (1)). Log sales is the log of one plus manager i's monthly sales credits as a worker, averaged over the 12 months prior to i's promotion (or for i's total pre-promotion tenure, if less than 12 months). Log collaborators is the log of one plus manager i's number of distinct collaborators each month, averaged over the 12 months prior to promotion (or for i's total pre-promotion tenure, if less than 12 months). Even-numbered columns include controls for the manager's tenure and team size in the month prior to promotion, as described in Table 3. Columns 3 and 4 further restrict the sample to managers who are assigned to subordinates that are not their previous teammates. Standard errors are adjusted for heteroskedasticity.

Table 6: Validity of Heckman Instruments

Manager value added	All positions			Promoted to different team		
	(1)	(2)	(3)	(4)	(5)	(6)
Industry-month promotion rate	-0.00289		-0.0177	0.000726		-0.0252
	(0.0183)		(0.0302)	(0.0187)		(0.0223)
Firm-month promotion rate		0.0134	0.00443		0.0174	0.0200
		(0.0126)	(0.0149)		(0.0130)	(0.0141)
R-squared	.000	.002	.014	.000	.004	.004
Observations	696	696	696	608	608	608

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table tests whether the industry-level and firm-level promotion rates predict our estimates of manager value added, within the sample of promoted managers. All variables and sample restrictions are as defined for Table 5. Standard errors are adjusted for heteroskedasticity.

Table 7: Heckman Selection Model: Predictors of Manager Value Added

Manager value added	All positions		Promoted to different team	
	(1)	(2)	(3)	(4)
Log(sales)	-0.103**	-0.202***	-0.122**	-0.207***
	(0.0498)	(0.0725)	(0.0503)	(0.0721)
Log(collaborators)	0.237**	0.156	0.223*	0.157
	(0.111)	(0.105)	(0.122)	(0.114)
Pre-promotion controls	Yes	Yes	Yes	Yes
Instrument	Promotion rate	Firm-month FE	Promotion rate	Firm-month FE
R-squared	.068	.092	.101	.12
Observations	696	696	608	608

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reestimates Table 5 using the Heckman selection model, using industry- and firm-level promotion rates or firm-year-month fixed effects as the instruments for selection into the sample of promoted managers. All other variables and sample restrictions are as defined for Table 5. Standard errors are adjusted for heteroskedasticity.

Table 8: Assignment of Managers to Subordinates

Log(subordinate sales), past months:	7-12	4-6	1-3	7-12	4-6	1-3
	(1)	(2)	(3)	(4)	(5)	(6)
Log(sales)	0.425***	0.442***	0.431***			
	(0.0479)	(0.0508)	(0.0522)			
Log(collaborators)	-0.00140	0.0274	-0.171			
	(0.105)	(0.129)	(0.126)			
Manager value added				-0.0127	0.0106	-0.0546
				(0.105)	(0.112)	(0.111)
R-squared	.230	.201	.184	.000	.000	.001
Observations	626	626	626	626	626	626

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table explores how new managers are assigned to subordinates. The sample is at the manager level, and includes all promoted managers for whom there exists data on manager and subordinate pre-promotion performance. The dependent variable is subordinate average monthly log sales in the 7-12 month, 4-6 month, and 1-3 month period prior to the promotion event. Log sales and log collaborators represent manager pre-promotion characteristics as defined in Table 5. Manager value added fixed effects are also as defined for Table 5. Standard errors are adjusted for heteroskedasticity.

Table 9: Splines: Non-linearities in Predictors of Promotion and Manager Value Added

	Worker is promoted		Manager value added	
	(1)	(2)	(3)	(4)
Log(sales) tercile 1	-0.00458		-0.180**	
	(0.00376)		(0.0729)	
Log(sales) tercile 2	0.0975***		-0.502*	
	(0.0115)		(0.267)	
Log(sales) tercile 3	0.0759***		-0.121	
	(0.0116)		(0.214)	
Log(sales) quintile 1		-0.00483		-0.144*
		(0.00447)		(0.0765)
Log(sales) quintile 2		0.00851		-0.530**
		(0.0159)		(0.256)
Log(sales) quintile 3		0.124***		-0.343
		(0.0224)		(0.532)
Log(sales) quintile 4		0.0821***		-0.334
		(0.0210)		(0.484)
Log(sales) quintile 5		0.0693***		-0.153
		(0.0181)		(0.253)
Pre-promotion characteristics	Yes	Yes	Yes	Yes
Firm-month FE	Yes	Yes	N/A	N/A
Instrument	N/A	N/A	Firm-month FE	Firm-month FE
R-squared	.056	.056	.095	.097
Observations	734278	734278	696	696

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Columns 1 and 2 examine non-linearities in the relation between promotion and worker sales performance. The specification is similar to that in Table 3 Column 4, except that we allow for a three part or five part spline with respect to Log(sales). The knots for the spline are spaced to evenly divide the sample into terciles or quintiles. The reported coefficients represent the slope within each interval (they do not represent marginal differences between intervals). Columns 3 and 4 examine non-linearities in the relation between manager value added and worker sales performance. The specification is similar to that in Table 7 Column 2, except that we allow for a three part or five part spline with respect to Log(sales). The knots for the spline are spaced to evenly divide the sample into terciles or quintiles. The reported coefficients represent the slope within each interval. Coefficients in Columns 1 and 2 have been multiplied by 100 to improve readability and represent the percentage point change in promotion probability for a unit change in each independent variable. Standard errors are clustered by worker in Columns 1 and 2 and adjusted for heteroskedasticity in Columns 3 and 4.

Table 10: Heterogeneity in Promotion Policies

	Worker is promoted	
	(1)	(2)
Log(sales)	0.0333***	0.0686***
	(0.00692)	(0.0142)
Log(sales) x fraction variable pay	-0.0450***	-0.104***
	(0.0143)	(0.0313)
Log(collaborators)	-0.0247	-0.0471
	(0.0159)	(0.0315)
Log(collaborators) x fraction variable pay	0.0701**	0.133*
	(0.0346)	(0.0690)
Fraction variable pay	0.449***	
	(0.170)	
Industry-month promotion rate	0.00285	
	(0.0637)	
Firm-month promotion rate	0.998***	
	(0.0561)	
Pre-promotion characteristics	Yes	Yes
Firm-month FE	No	Yes
R-squared	.048	.048
Observations	473904	473904

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table explores how promotion policies vary with the strength of pay for performance across firms. Fraction variable pay represents the ratio of commissions plus bonus to salary, averaged across all workers within each firm-year. Observations corresponding to firms for which we do not have compensation data are excluded from the sample. All other variables and sample restrictions are as described in Table 3. Standard errors are allowed to be clustered by worker.