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

Alan Benson, Danielle Li, and Kelly Shue*

September 1, 2016

Abstract

The best worker isn't 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 quality. We estimate that the costs of managerial mismatch are substantial, suggesting that firms make inefficient promotion decisions or 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, 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, and Wharton People Analytics Conference for helpful comments. We thank Menaka Hamplole and Leland Bybee for excellent research assistance. Correspondence: bensona@umn.edu.

The best workers do not always make the best managers. 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 subordinates, 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 classic settings for the Peter Principle (such as promoting excellent engineers into management), the performance of sales workers is also readily measurable.

Our analysis uses new data that are uniquely well suited to examine whether firms' promotion policies maximize managerial quality. These data, provided by a firm that offers sales performance management software over the cloud, include standardized measures of sales transactions, organizational hierarchy, and compensation for a panel of 48,209 workers matched to 5,369 managers at 214 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 the 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 characteristics. For promoted managers, we calculate their "value-added" in shaping their subordinates' sales performance, i.e., each manager's contribution to improving their subordinates' sales.

We begin by showing that past sales performance is a strong positive predictor of promotion.

[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."

Yet, among observed promotions, pre-promotion sales performance is in fact negatively correlated with managerial performance: on average, doubling a new manager's pre-promotion sales performance 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 suggests that firms underweight sales collaboration experience in promotion decisions. Pre-promotion collaboration experience is a strong positive predictor of managerial performance. However, workers with more collaboration experience are not more likely to be promoted after we account for firm-level time trends.

Overall, we find empirical support for the Peter Principle: firms promote based upon current job performance even though pre-promotion sales negatively predicts managerial performance and other observable characteristics positively predict managerial performance. However, we face the empirical challenge that we only observe measures of managerial performance among workers who are actually promoted. Ideally, we would observe latent managerial potential for all workers and test whether firms promote the workers with the greatest managerial potential. Given that we only observe managerial performance among the promoted, it remains possible that managers with high pre-promotion sales are more likely to have "just crossed" the promotion threshold. If so, we would observe a negative relation between pre-promotion sales and managerial performance for the sample of promoted workers, even though firms do promote the best potential managers.

To address this measurement challenge, we apply a model of promotions based on the Heckman (1979) selection model with refinements by Chandra and Staiger (2011, 2016) and Abaluck and Agha (2014). In the model, firms hold beliefs about workers' managerial potential based upon worker sales performance, collaboration experience, tenure, and other observables. Separately, firms condition promotion on these same characteristics. The model allows us to measure whether firms promote the best potential managers but does not impose any normative judgments regarding

what firms should do. Firms may not promote the best potential managers for two reasons: firms are mistaken about how observable worker characteristics predict managerial potential or firms purposely "lower the bar" for workers with certain characteristics to achieve other objectives such as improved incentives.

Estimating the model involves a two step procedure. First, using the sample of all workers, we estimate each workers' propensity to be promoted as a function of both their sales performance and other observable characteristics. Second, using the sample of promoted workers, we compare the performance of managers with the same propensity to be promoted. If firms overweight past sales performance in promotion decisions, then past sales performance will be negatively correlated with managerial performance among managers with an equal propensity to be promoted. Intuitively, if firms are promoting to maximize match quality, then they will only care about sales performance and other worker observables insofar as they impact managerial performance and should adjust a worker's propensity to be promoted accordingly. Controlling for promotion propensity, better salespeople should not make systematically worse managers. If this were the case, then firms are prioritizing prior sales performance in promotion decisions at the expense of maximizing managerial performance.

We find that promotion decisions overweight the past sales performance of workers. Holding a worker's propensity to be promoted constant, better salespeople make worse managers. Further, firms could improve managerial performance by placing more weight on worker collaboration experience. Overall, we estimate that firms could improve total sales by the equivalent of one extra salesperson per five-person team if the allocation of promotions could be changed without any of its potential costs.

We caution that our results do not necessarily imply that firms use suboptimal promotion policies. 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 upon 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 discouraging the manipulation of other, more fungible performance metrics

such as credit sharing (DeVaro and Gürtler, 2015). While our analysis doesn't necessarily imply suboptimal promotion policies, our results do suggest that the costs of not promoting the best potential managers are high.

We also explore two potential alternative explanations of our results. First, Lazear (2004) shows that mean reversion can create the illusion of the Peter Principle. Specifically, Lazear notes that high performing salespeople may indeed be the best predicted managers, but their performance would nonetheless decline following promotion if their own performance were mean-reverting. Our results cannot be explained by mean reversion because we measure managerial performance not by the manager's own sales performance (which may be mean-reverting), but by the improvement of the manager's subordinates. Further, 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 relationship between pre-promotion sales and post-promotion managerial performance, as we find in the data. Second, we show that our results are unlikely to be explained by differences in the way promoted workers are matched to sales teams. Promoted workers with strong previous sales performance are matched to better teams of subordinates and these teams do not appear to be on differential growth trends.

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, 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.

Finally, our findings 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 upon 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. Salespeople 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.

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.

[Table 1]

1.1 Overview of sales positions

Sales workers are typically assigned a market consisting of a territory, 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 skew 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 slightly skewed normal distribution. The bottom left panel, which reflects our measure of sales performance, shows the residual distribution of monthly sales after controlling for firm-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 company-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 company-month, the correlation between our sales and collaboration measures is 0.28. While positive and statistically significant, the moderately small 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.[2]

Note that 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 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.

[2]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.

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 manager.

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 are 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 6, presented later in the text, 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 (2015), as well as a large literature on employer-employee, and teacher-student matched data (e.g., Abowd, Kramarz, and Margolis, 2002), and estimate a manager's performance as his or her *value-added*, controlling for each worker's average performance across different managers. 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 outcome of interest is the log of worker i's sales performance under manager m in firm f in month t. δ_i, δ_m, and $\delta_{f\times t}$ represent worker, manager, and firm-month fixed effects, respectively.[3] 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.

Our estimates may still suffer from bias if managers are non-randomly assigned to subordinates

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

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 attribute this improvement to the managers. Yet even if this were the case, these 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 4.5, 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 lumpy, with many workers making zero sales in some months and large sales in others.[4] For example, a worker may be observed under one manager, A, in a month in which he happens not to sell any products, and under a different manager, B, in a month where he happens to make a large sale. For this specific worker, it may mistakenly look as if manager B improved his performance. As long as workers are not systematically more likely to be assigned to manager B in months when they would have made more sales anyway, this should not bias our estimates of manager fixed effects. In particular, our estimates of the Peter Principle, i.e., the extent to which manager pre-promotion characteristics predict mean differences in manager value added remain unbiased. However, this additional noise does lead to an upward bias in our estimates of the *variance* of manager fixed effects.

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 (2,569 do) 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, his new

[4]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.

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 value added and pre-promotion performance for 654 managers.

[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 managerial fixed effects, which exaggerates the variance. Nevertheless, our later regression estimates of the mean expected differences in manager value added across managers with different pre-promotion attributes will remain unbiased.

2 What predicts promotion?

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

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

We estimate Equation (2) on a worker-month level panel for workers i at firm f who have not yet been promoted as of 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 X_{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). X_{ift} also includes fixed effects for seven bins of worker tenure, interacted with whether tenure may be censored in the data.

[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 company-month fixed effects, Column 2 shows that a 10 percent improvement in sales credits increases the likelihood that a worker is promoted by 0.32 percentage points, relative to a base rate probability of being promoted each month of 0.21 percent. This amounts to more than a doubling in promotion probability. Our results also imply that a doubling in sales corresponds to a more than ten-fold increase in promotion probability. By contrast, we do not find consistent evidence that firms weigh collaboration experience in promotion decisions. Once we control for company-month fixed effects, the correlation between collaboration and promotion is both economically small and statistically insignificant.

[Figure 3]

This pattern is underscored by Figure 3, which plots the distribution of workers' fitted propensity 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 2 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 company-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 2 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 sales experience. Controlling for sales, 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:

$$\begin{aligned} \text{Manager Value Added}_{if} &= a_1 \text{Pre-Promotion Sales}_{if} + a_2 \text{Pre-Promotion Collaborators}_{if} \\ &\quad + \delta_f + u_{if} \end{aligned} \tag{3}$$

We estimate Equation (3) at the manager level. 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, $\text{Collaborators}_{if}$ 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). To more efficiently estimate the firm fixed effects, the sample includes both managers who have been promoted, for whom we observe pre-promotion characteristics, as well as those who have always been managers during our sample period. For those who have always been managers in our sample, Pre-Promotion Sales_{if} and Pre-Promotion $\text{Collaborators}_{if}$ are missing and set to zero, and we introduce an indicator for a non-promoted manager as an additional control variable.

[Table 4]

Table 4 shows that, among promoted managers, there is a significant *negative* relation between pre-promotion sales performance and subsequent managerial performance. A 10 percent increase in a manager's pre-promotion sales corresponds to an approximate 0.9 percentage point decline in manager value added. Since manager value added represents the change in log subordinate sales, this implies that a manager with slightly higher pre-promotion sales of 10 percent leads each subordinate's sales to decline by almost 1 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 implies that total team sales under the new manager will decline by approximately as much as if the team lost half of one worker. In contrast, collaboration experience is positively correlated with manager value added. In Column 2 with company fixed effects, we find that a 10 percent increase in collaboration experience is correlated with a large 2.7 percentage point improvement in manager value added.

In some cases, workers are promoted to replace their former managers. To ensure that these observations are not driving our results, we can further restrict the 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 worker prior to promotion. In Columns 3 and 4 of Table 4, 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.

4 Are firms promoting the best potential managers?

The empirical results so far show that firms promote based upon 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 only observe measures of managerial performance among workers

who are actually promoted. In an ideal test of whether firms promote the best potential managers, we would observe latent managerial potential for all workers and test whether firms promote those with the greatest managerial potential. Given that we only observe managerial performance among the promoted, it remains possible that managers with high pre-promotion sales are more likely to have "just crossed" the promotion threshold. If so, we may observe a negative relation between pre-promotion sales and managerial performance for the sample of promoted workers, even though firms do promote the best potential managers.

We next discuss a simple model of firm promotions and illustrate how it can be used to assess whether firms are indeed promoting the best potential managers.

4.1 Theory

Firms may value sales performance in promotion decisions for three reasons: (1) workers with high sales performance make better managers, (2) promoting workers with higher sales performance is associated with other benefits, such as improved incentives for workers to exert effort, increases in morale, or preservation of fairness concerns, or (3) firms mistakenly overweight the value of sales performance in predicting managerial performance. For brevity, we respectively refer to these as the matching, incentive, and biased-belief motivations. In this section, we develop an empirical framework that allows us to decompose observed promotion patterns into a portion that is driven by matching, and another portion that is driven by incentive motives or biased beliefs. We are unable to distinguish between the incentive and biased belief motivations, but by separating out the extent to which firms do not promote the best potential managers, we provide a lower bound for the incentive benefits that firms must expect in order to rationalize their promotion decisions.

Our approach to modeling promotions is based on classic selection models developed by Heckman (1979) and extended by Chandra and Staiger (2010) and Abaluck and Agha (2015) to study bias in the allocation of medical treatment to minority groups, where the econometrician only observes the effect of the medical treatment on the treated. The previous studies are able to identify prejudice in medical treatment against minorities by showing that minority patients who do receive medical treatment benefit more than other patients with equal propensity to be treated. The underlying

Beckerian intuition of these models is that prejudice against minority patients imply that minority patients will have higher marginal returns from being treated, holding propensity to treat constant. In our setting, if firms "discriminate" against non-top salespeople in promotion decisions, then the observed non-top salespeople who are promoted should perform better than those with higher sales, holding propensity to promote constant.

In our model, firms make promotion decisions based on worker observable characteristics. Empirically, we focus on two measures of worker level performance: sales performance (Sales_i) and collaboration experience (Collaborators_i). Other observable worker characteristics include firm affiliation, team size, and tenure (X_i), described in more detail in Section 4.2. Worker characteristics that are observed by firms but not by the econometrician are captured by ε_i.

Suppose that the true relationship between these characteristics and a worker's latent managerial ability M_i is given by:

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

Firms may systematically misweight the importance of a worker's sales and collaboration experience, so that a firm's beliefs about the worker's potential managerial performance are given by:

$$M_i' = (\beta_1 + \beta_1^e)\text{Sales}_i + (\beta_2 + \beta_2^e)\text{Collaborators}_i + X_i\beta_3 + \varepsilon_i \tag{5}$$

where β_1^e and β_2^e denote errors in the firm's beliefs. The model assumes that firms have correct overall beliefs about the other worker characteristics X_i and unobserved characteristics ε_i. We discuss the reasons for this assumption in Section 4.3.

Firms promote workers based on their beliefs about workers' potential managerial performance M_i', but may apply different acceptable thresholds τ_i at which they would be willing to promote any given worker. Firms promote if $M_i' > \tau_i$, where the threshold τ_i may depend on worker sales or collaboration experience:

$$\tau_i = \tau_0 + \tau_1 \text{Sales}_i + \tau_2 \text{Collaborators}_i \tag{6}$$

Equation (6) allows firms to prioritize sales performance in promotion decisions, even for salespeople with identical managerial potential. For example, firms that want to incentivize high sales could adopt a lower promotion threshold for salespeople with higher sales, setting $\tau_1 < 0$. In this case, high performing salespeople could get promoted ahead of lower-performing salespeople who are expected to be better managers—a practice consistent with trading match quality for stronger incentives. If the firm wants to incentivize collaboration, it may instead adopt $\tau_2 < 0$. Alternatively, if firms only want to promote the best potential managers, then they would retain the same threshold for all salespeople, such that $\tau_1 = \tau_2 = 0$.

Together, Equations (5) and (6) imply a worker's probability of promotion is:

$$\begin{aligned}
\Pr(\text{Promote}_i) &= \Pr(M_i > \tau_i) \\
&= \Pr((\beta_1 + \beta_1^e)\text{Sales}_i + (\beta_2 + \beta_2^e)\text{Collaborators}_i \\
&\quad + X_i\beta_3 + \varepsilon_i > \tau_0 + \tau_1\text{Sales}_i + \tau_2\text{Collaborators}_i) \\
&= \Pr(I_i > -\varepsilon_i) \qquad (7) \\
&\quad \text{where } I_i \equiv (\beta_1 + \beta_1^e - \tau_1)\text{Sales}_i + (\beta_2 + \beta_2^e - \tau_2)\text{Collaborators}_i + X_i\beta_3 - \tau_0
\end{aligned}$$

In Equation (7), $Pr(I_i > -\varepsilon_i)$ denotes worker i's propensity to be promoted and two workers with the same I will have the same propensity to be promoted. Intuitively, a worker with high sales is more likely to be promoted for two reasons. The first is that better salespeople tend to make better managers (if $\beta_1 > 0$). In this case, firms promote high-performing salespeople in order to maximize match quality. The second is that firms apply a lower promotion threshold for good sales workers or have positively biased beliefs regarding the managerial potential of high performing salespeople ($\beta_1^e - \tau_1 > 0$). In particular, $\tau_1 < 0$ would be consistent with cases in which firms forgo managerial match quality to prioritize the chief pre-promotion performance metric for incentive reasons. Whether due to mistaken beliefs or incentive concerns, we define ($\beta_1^e - \tau_1 > 0$) or equivalently ($\tau_1 - \beta_1^e < 0$) to be the hallmark of the Peter Principle.

To identify the sign of $\tau_1 - \beta_1^e$, we examine the expected managerial performance among

promoted workers, i.e., the treatment on the treated effect:

$$
\begin{aligned}
E(M_i|\text{Promote}) =& E(M_i|I_i > -\varepsilon_i) \\
=& \beta_1 \text{Sales}_i + \beta_2 \text{Collaborators}_i + X\beta_3 + E(\varepsilon|I_i > -\varepsilon_i) \\
=& (\beta_1 + \beta_1^e - \tau_1)\text{Sales}_i + (\beta_2 + \beta_2^e - \tau_2)\text{Collaborators}_i + X_i\beta_3 - \tau_0 \\
& + (\tau_1 - \beta_1^e)\text{Sales}_i + (\tau_2 - \beta_2^e)\text{Collaborators}_i + E(\varepsilon_i|I_i > -\varepsilon_i) \\
=& I_i + (\tau_1 - \beta_1^e)\text{Sales}_i + (\tau_2 - \beta_2^e)\text{Collaborators}_i + \tau_0 + E(\varepsilon|I_i > -\varepsilon) \\
=& (\tau_1 - \beta_1^e)\text{Sales}_i + (\tau_2 - \beta_2^e)\text{Collaborators}_i + g(I_i) \\
& \text{where } g(I_i) \equiv \tau_0 + I_i + E(\varepsilon|I_i > -\varepsilon)
\end{aligned} \tag{8}
$$

The key insight offered by Equation (8) is that we can identify $(\tau_1 - \beta_1^e)$ and $(\tau_2 - \beta_2^e)$ by controlling flexibly for I, which determines a worker's propensity to be promoted. Holding $g(I)$ constant, Equation (8) compares the managerial performance of two promoted workers who have the same likelihood of being promoted. If the firm has unbiased beliefs and applies the same promotion threshold to everyone, then after controlling for the propensity to promote, higher performing salespeople should not be systematically better or worse managers. In other words, if firms only seek to promote the best potential managers, then all factors that predict managerial potential should enter through $g(I)$. Any excess correlation ($\tau_1 - \beta_1^e \neq 0$ or $\tau_2 - \beta_2^e \neq 0$) indicates that the firm either has mistaken beliefs or purposely conditions promotion thresholds on workers' sales and collaboration experience.

Equation (8) also highlights why it is important to control for each worker's promotion propensity. Consider a simpler regression of managerial performance regressed on sales performance and collaboration experience, estimated using the sample of promoted workers:

$$
M_i = a_0 + a_1 \text{Sales}_i + a_2 \text{Collaborators}_i + e_i \tag{9}
$$

Suppose that the estimated coefficient a_1 is negative, implying that, among promoted workers, better salespeople tend to be worse managers. This fact is consistent with the Peter Principle:

firms apply a lower promotion threshold for high performing workers ($\tau_1 < 0$) or mistakenly believe that pre-promotion sales is more positively predictive of managerial potential than it is ($\beta_1^e > 0$). However, another potential explanation is that firms apply the same promotion threshold to all workers and have correct beliefs, but the distribution of promotion propensity differs by sale performance, such that managers with high pre-promotion sales are more likely to have "just crossed" the promotion threshold. If high sales workers that are promoted on average have lower propensity to be promoted, then the estimated a_1 would be negative, even if firms promote the best potential managers.

4.2 Estimation strategy

We measure promotion propensity by estimating the regression in Equation (2) on the full worker-month level panel. Using the estimated coefficients, we then construct fitted values representing I_{ift} for each worker-firm-month. For each worker who is actually promoted, we use the worker's predicted I in the month prior to promotion. Next, we control flexibly for promotion propensity using a quintic with respect to I_i, in a regression of realized managerial performance on pre-promotion characteristics:

$$\begin{aligned} \text{Manager Value Added}_{if} &= b_1 \text{Pre-Promotion Sales}_{if} + b_2 \text{Pre-Promotion Collaborators}_{if} \\ &\quad + I_m + \cdots + I_m^5 + \delta_f + u_{if} \end{aligned} \tag{10}$$

We estimate this regression at the manager level for the sample of promoted managers. Equation (10) is the regression analogue of Equation (7) from our model. It is also the same as Equation (3), but augmented with a flexible set of controls $I_i + \cdots + I_i^5$ for manager i's promotion propensity as a worker.

b_1 and b_2 are the primary coefficients on interest, and correspond to estimates of $\tau_1 - \beta_1^e$ and $\tau_2 - \beta_2^e$ in Equation (7). The null hypothesis is that firms maximize managerial match quality, implying $b_1 = b_2 = 0$. In other words, controlling for the propensity to promote, salespeople with

greater sales performance or more collaboration experience should not be systematically better or worse managers. Any excess correlation captured by $b_1 \neq 0$ or $b_2 \neq 0$ indicates that the firm either has mistaken beliefs or conditions promotion thresholds on sales and collaboration experience. If we find $b_1 < 0$, this indicates that firms could improve managerial match quality by placing less weight on sales performance in their promotion decisions. The same logic applies for b_2 and collaboration experience.

4.3 Identifying assumptions and interpretation

As with standard Heckman selection models, the estimation of Equation (10) relies either on functional form restrictions on $g(I)$ or exclusion restrictions. If we include the same set of control variables in the first stage promotion equation (2) as in the second stage managerial performance regression (10), then the estimated promotion propensity from the first stage will be a linear combination of the covariates in the second stage regression. In this case, the coefficients b_1 and b_2 will be identified only from the functional form imposed on $g(I)$: in our case, the quintic polynomials in I_m. To avoid this, we exclude the variables X_{ift} from entering the second stage equation except through $g(I)$. Because these variables are excluded, we can think of them as instrumental variables that aid the estimation of I, but which do not impact manager value added except through the promotion process described by I. This reflects a standard IV identification in Heckman selection models.

Substantively, this exclusion restriction requires that firms do not hold biased beliefs regarding how the variables included in X_{ift} (e.g., tenure and team size) predict managerial potential and also do not set differential promotion thresholds with respect to these variables. That is, we allow firms to misweight sales performance and collaboration experience in predicting managerial performance (i.e., $\tau_1 - \beta_1^e$ and $\tau_2 - \beta_2^e$ need not be zero), but we assume that firms do not misweight these other variables. This restriction is equivalent to including these additional covariates directly in the second stage regression, but imposing that the coefficient on these variables be equal to zero after controlling for $g(I)$.

In practice, the Peter Principle could manifest through firms misweighting other measures

of workers' current performance that are imperfectly captured by sales credits (which is the most complete measure of performance available in our data sample). If so, these other worker characteristics should separately enter Equation (10), but we would mistakenly omit them. To the extent that these other characteristics and sales performance are correlated, our estimate of b_1 may reflect the firm's misweighting of other worker characteristics that are correlated with sales credits. The same logic applies to omitted variables that are correlated with collaboration experience.

We view violations of this exclusion restriction as changing the interpretation of our results, rather than invalidating the exercise. Under the null hypothesis that firms promote the best potential managers, any worker characteristic, including our sales and collaboration measures, should have a zero coefficient in the second stage regression after controlling for the worker's promotion propensity. As such, finding either $b_1 \neq 0$ or $b_2 \neq 0$ is evidence against this null and in favor of the fact that firms are not maximizing managerial match quality, and that they instead prioritize or deemphasize something correlated with sales performance or collaboration experience. In our counterfactual simulations, we can still examine the magnitude of the match quality loss associated with these promotion policies.

Our identification strategy also relies on a single-index selectivity assumption that is common to many Heckman-style selection models. We assume that $g(I)$ depends only on the index, i.e., that $E(\varepsilon_i|I_i > -\varepsilon_i)$ can be expressed as a function of I_i only. This is equivalent to requiring that the distribution of ε_i (worker characteristics that predict managerial performance and are observable to the firm but unobserved to the econometrician) does not differ by sales performance or collaboration experience. If the variation in ε_i is greater for high sales workers, then the truncated mean $E(\varepsilon_i|I_i > -\varepsilon_i)$ would be larger for any given truncation point, resulting in higher expected managerial performance for high sales workers conditional on promotion. This could occur if, for example, firms groom high performers for management positions, and in doing so, they receive more information about worker characteristics captured in ε_i. This would be a bias against our findings in favor of the Peter Principle. Alternatively, if the variation in ε_i is smaller for high sales workers, then $E(\varepsilon_i|I_i > -\varepsilon_i)$ may be smaller among promoted workers with high sales, leading to lower expected managerial performance conditional on promotion. This would appear as if firms

were favoring high sales workers in promotion decisions even if they were not.

4.4 Results

Table 5 shows that among promoted workers with the same *ex ante* propensity to be promoted, those with greater pre-promotion sales systematically become worse managers. To give a sense of the magnitude of this effect, consider two managers at the same firm with the same propensity to be promoted but different pre-promotion performance records. Our results imply that a 10 percent improvement in pre-promotion sales is associated with 0.9 percent decline in sales for each subordinate (Column 2). By contrast, controlling for promotion propensity, pre-promotion collaboration experience is positively correlated with managerial performance: a 10 percent improvement in collaboration experience is associated with a 2.6 percent increase in sales per subordinate. These results are robust to controlling for company fixed effects.[5]

[Table 5]

In columns 5 and 6, 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 4]

These results suggest that firms "discriminate" in favor of strong salespeople by lowering their promotion thresholds or holding biased beliefs about the relationship between sales and managerial quality. An implication of this behavior is that, among promoted managers with the same propensity to be promoted, weaker salespeople should be more successful. Figure 4 provides direct evidence for this prediction. The top panel plots the relationship between residual

[5] In Table 5, we include only measures of sales performance, collaboration experience, and promotion propensity in our second stage estimation of Equation (10). As discussed earlier, we implicitly assume that firms accurately weigh other pre-promotion variables. A caveat for the interpretation of the results is that if we find that firms overweight sales performance, firms may be directly overweighting sales performance, or overweighting another variable that is correlated with sales performance.

managerial performance (after controlling for company fixed effects), among promoted managers for the same promotion propensity. The blue line represents managers from the bottom half of pre-promotion sales distribution while the red line plots the same relationship for managers from the top half of the pre-promotion sales distribution, again 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.

The overall slope of the relationship between promotion propensity and managerial performance also provides evidence in support of this explanation. Among high performing sales workers, there is no relationship between promotion propensity and managerial performance, suggesting that firms care about sales performance in itself when making promotion decisions, and not its relationship to managerial performance. Among lower performing sales workers, however, there is a positive relationship between promotion propensity and realized managerial performance, suggesting that firms only promote workers with poor sales if firms have reason to believe that these workers will make good managers.

The bottom panel of Figure 4 illustrates this same relationship, except with managers separated by their pre-promotion collaboration experience. Here we see the opposite pattern. Conditional on promotion propensity, managers who collaborated with more people prior to promotion consistently have higher value added.

4.5 Alternative explanations and robustness

The results above suggest that firms are promoting high-performing sales workers at the expense of promoting the best potential managers. However, one may be concerned that there are other potential explanations for these facts.

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 manager performance is not based on a manager's own sales, but is instead based on 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 does not predict that, across individuals, pre-promotion sales should be *negatively* correlated with post-promotion managerial performance, as we find in the data. In our sample, firms could promote better managers by simply selecting workers is lower sales performance, all else equal.

A second 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 managerial 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 6 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 a 2.3 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 assigned to subordinates with different time trends in performance. Table 6 columns 1, 3, and 5 examine subordinates' sales 12-months, 6-months, and 3-months prior to the arrival of the new managers, and finds no evidence of either a positive or negative time trend.

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 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 6 columns 2, 4, and 6 show that this is not the case. Instead, we find a statistically insignificant and inconsistently signed relationship between subordinate prior sales and manager value added.

One may also be concerned that some top sales workers may prefer not to be promoted. Although most workers enjoy significant pay increases after promotion, the very top sales workers can earn more than average sales managers. 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 given their managerial potential with their pay as sales workers, and then decide whether to accept the promotion. Thus, the workers most likely to accept promotions are those with comparatively good prospects as managers, so we should expect the greatest positive selection into management among the very best sales workers. 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 may actually make good managers, but prefer to remain in their current roles, our results still indicate that firms are not maximizing managerial quality 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.

Tables in the Appendix show results from several robustness tests. In Table 5, we use information on managers with value added estimates but without pre-promotion performance information to estimate company fixed effects with greater precision. Appendix Table A confirms that our results are robust to restricting our main sample to the 654 managers for whom we observe both value added and performance prior to promotion. Appendix Table B shows that our results are not driven by outlier observations by using estimates of manager value added based on worker sales variables winsorized at the two percent level on both sides.

4.6 What are the performance losses from mismatch?

How much do suboptimal promotions cost firms? 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 mismatch. Our estimates may be interpreted as the match quality that firms forgo to achieve better incentives.

Specifically, we examine how predicted managerial quality differs among three categories of workers: (1) actual promoted salespeople, (2) non-promoted salespeople, and (3) the top predicted manager among a promoted sales worker's peers, with peers defined as other salespeople in a team managed by the same manager. We interpret the last case to be the optimal promotion 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 increase.

[Figure 5]

Figure 5 shows the distributions of predicted manager value added in these three samples. To provide an intuition of the magnitudes, the median predicted improvement in subordinate salesperson performance is 2.6 percent for the promoted sample and -1.8 percent for the nonpromoted sample, consistent with firms doing slightly better than promoting at random. The median predicted improvement under the optimal promotion policy is 21 percent. Given that the typical manager is in charge of five subordinates, we estimate that an optimal promotion policy would lead to performance improvements roughly equivalent to one additional sales worker per team.

5 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 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 are behaving 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. 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.

6 Works Cited

Abaluck, Jason, Agha, Leila, Kabrhel, Christopher, Raja, Ali, and Venkatesh, Arjun. 2016. "Negative tests and the efficiency of medical care: what determines heterogeneity in imaging behavior?" forthcoming American Economic Review.

Baker, George, Gibbs, Michael, and Holmstrom, Bengt, 1994. "The wage policy of a firm." *The Quarterly Journal of Economics*, 921-955.

Baker, G.P., Jensen, M.C., Murphy, K.J., 1988. "Compensation and incentives: Practice vs. theory." *The Journal of Finance*, 43: 593-616.

Chandra, Amitabh, and Staiger, Douglas, 2010. "Identifying provider prejudice in healthcare." National Bureau of Economic Research working paper.

Chandra, Amitabh, and Staiger, Douglas. 2016. "Sources of inefficiency in healthcare and education." *American Economic Review: Papers and Proceedings*, 106(5): 383-7.

Deutsch, Claudia. 1986. "Holding on to technical talent." New York Times, 11/16/16.

DeVaro, Jed, 2006. "Internal promotion competitions in firms." *The Rand Journal of Economics*, 37: 521542.

DeVaro, Jed, and Gürtler , Oliver. 2015. "Strategic shirking in promotion tournaments." *The Journal of Law, Economics, and Organization.*

Fairburn, James A., and Malcomson, James M., 2001. "Performance, promotion, and the Peter Principle." *The Review of Economic Studies*, 68: 45-66.

Grabner, Isabella, and Moers, Frank. 2013. "Managers' choices of performance measures in promotion decisions: An analysis of alternative job assignments." *Journal of Accounting Research*, 51(5): 1187-220.

Heckman, James. 1979. "Sample selection bias as a specification error." *Econometrica*, 47(1): 153-61.

Kaplan, Steven N., Klebanov, Mark M., and Sorensen Morten. 2012. "Which CEO characteristics and abilities matter?" *Journal of Finance*, 67(3): 9713-1007.

Kaplan, Steven N. and Sorensen, Morten. 2016. "Are CEOs different? Characteristics of top managers." Working Paper.

Lazear, Edward. 2004. "The Peter Principle: A Theory of Decline." *Journal of Political Economy*, 112: S141S163.

Lazear, Edward, Rosen, Sherwin, 1981. "Rank-Order Tournaments as Optimum Incentive Contracts." *Journal of Political Economy*, 89(5): 841-64.

Medoff, James and Abraham, Katherine. 1980. "Experience, performance, and earnings." National Bureau of Economic Research Working Paper.

Murphy, Kevin J., and Ján Zábojník. 2004. "CEO pay and appointments: A market-based explanation for recent trends." *American Economic Review: Papers and Proceedings*: 94(2): 192-6.

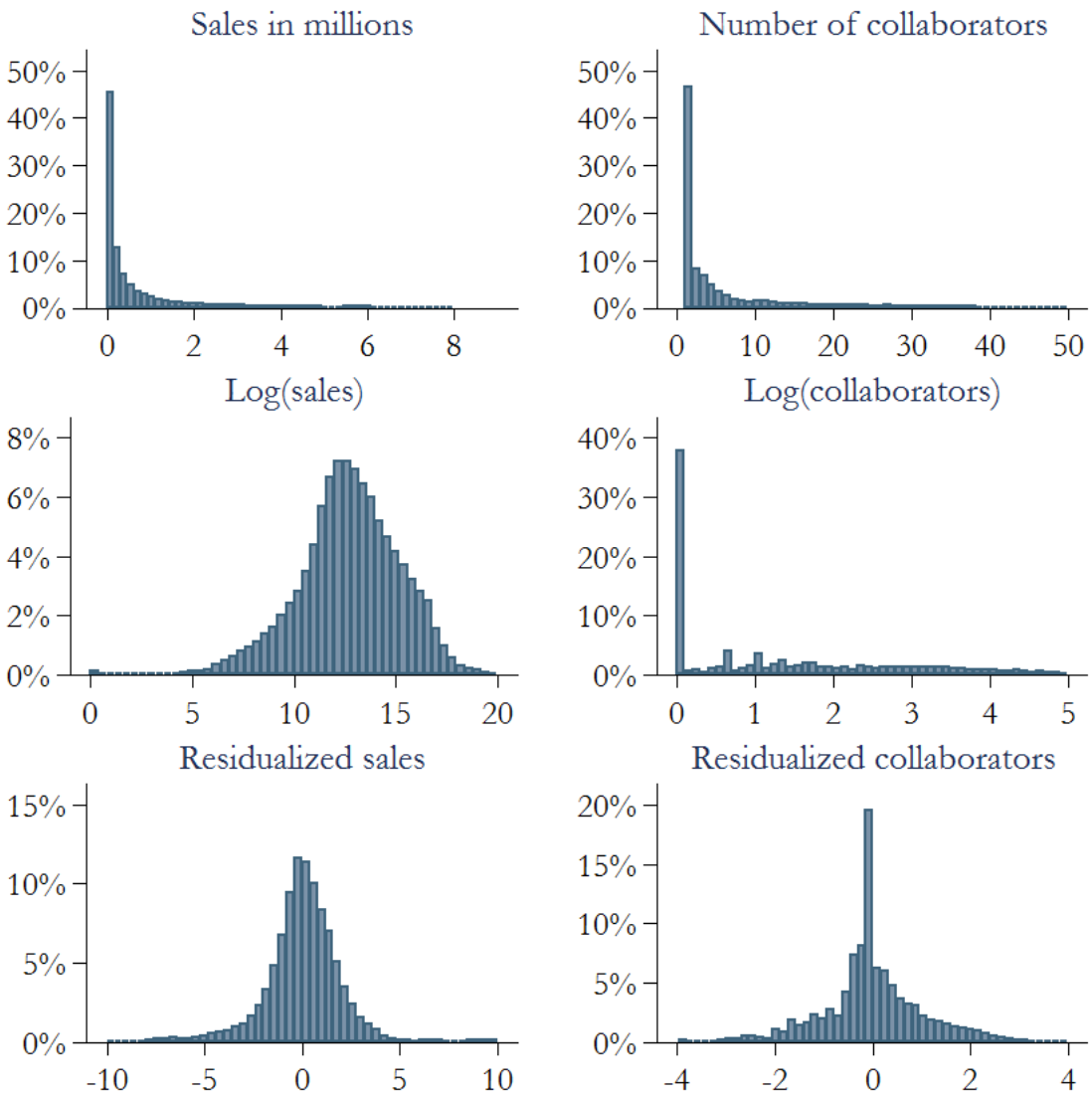
Peter, Laurence, and Hull, Raymond. 1969. *The Peter Principle*. William and Morrow and Co.

Prendergast, Canice. 1998. "What happens within firms? A survey of empirical evidence on compensation policies." *Labor statistics measurement issues*. University of Chicago Press.

U.S. Census Bureau. 2011. *Occupational Outlook Handbook*. Retrieved online from http://www.bls.gov/oes/2011/may/oes410000.htm.

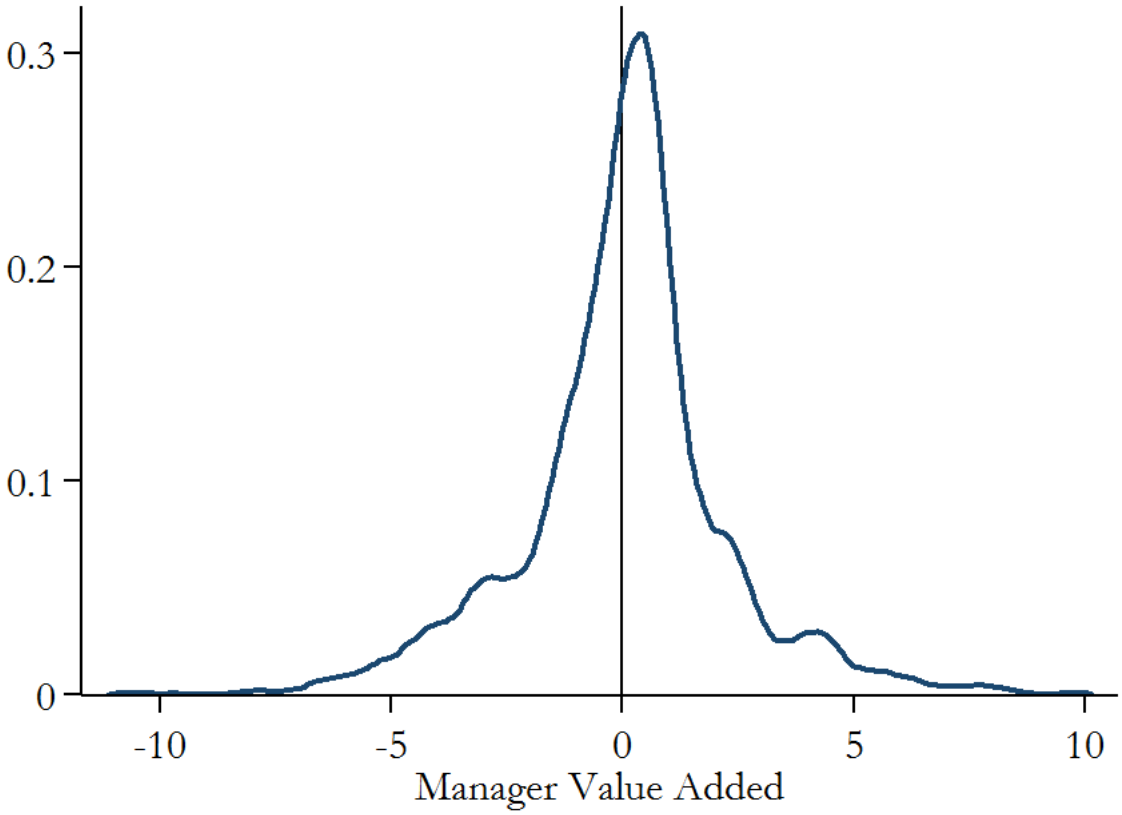
Waldman, Michael., 2003. "Ex ante versus ex post optimal promotion rules: The case of internal promotion." *Economic Inquiry*, 41: 27-41.

Figure 1: Distribution of sales and number of collaborators



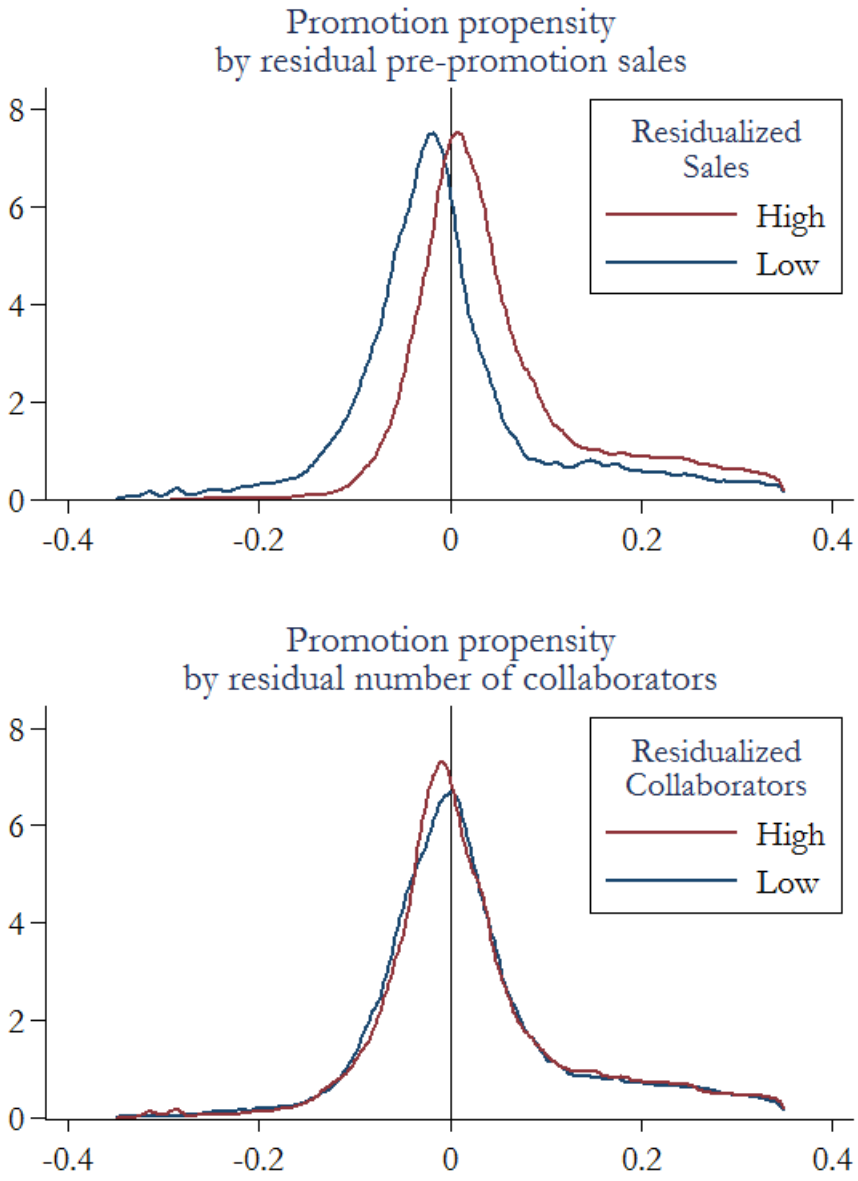
Notes: Left and right panels are 12-month moving averages for sales and number of collaborators, respectively. The top panels show the untransformed distribution. The middle panels show the log-transformed distribution. The bottom panels show the residuals after the log-transformed variables are regressed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

Figure 2: Distribution of Manager Value Added



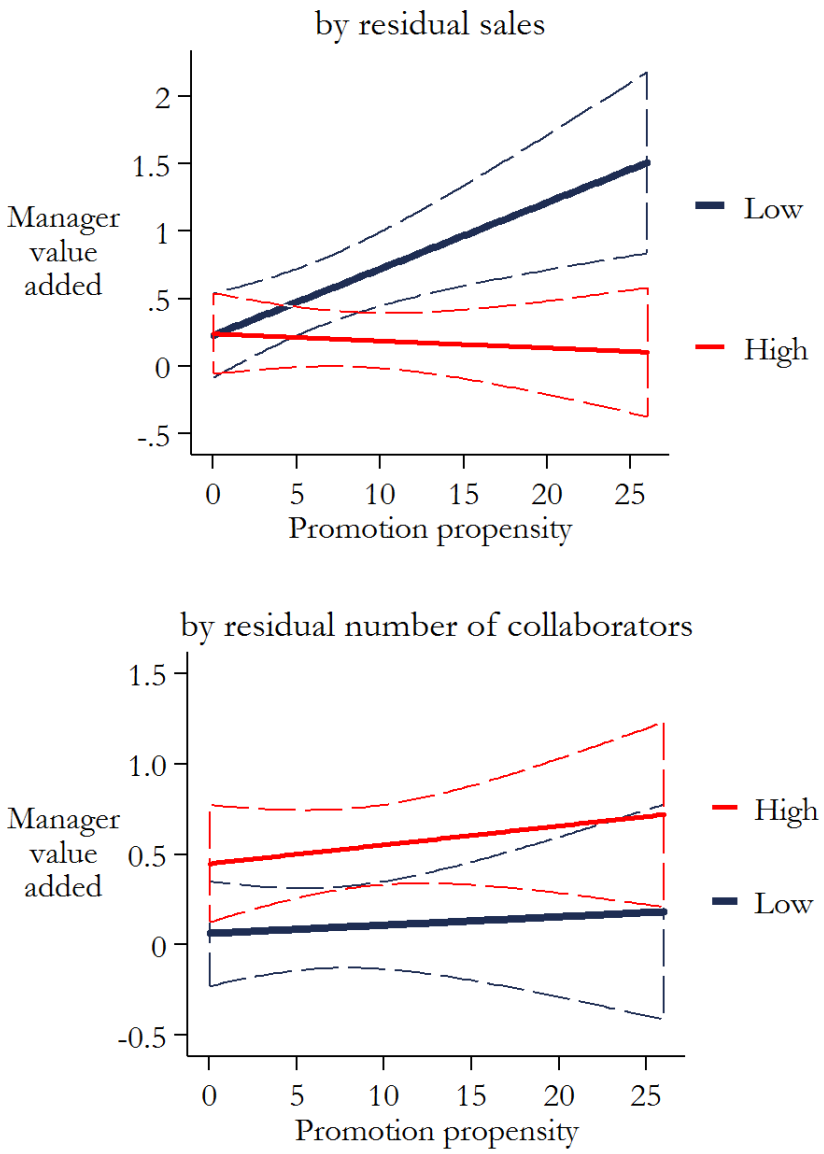
Notes: The figure shows the kernel density of estimated manager value added.

Figure 3: Distribution of Promotion Propensity



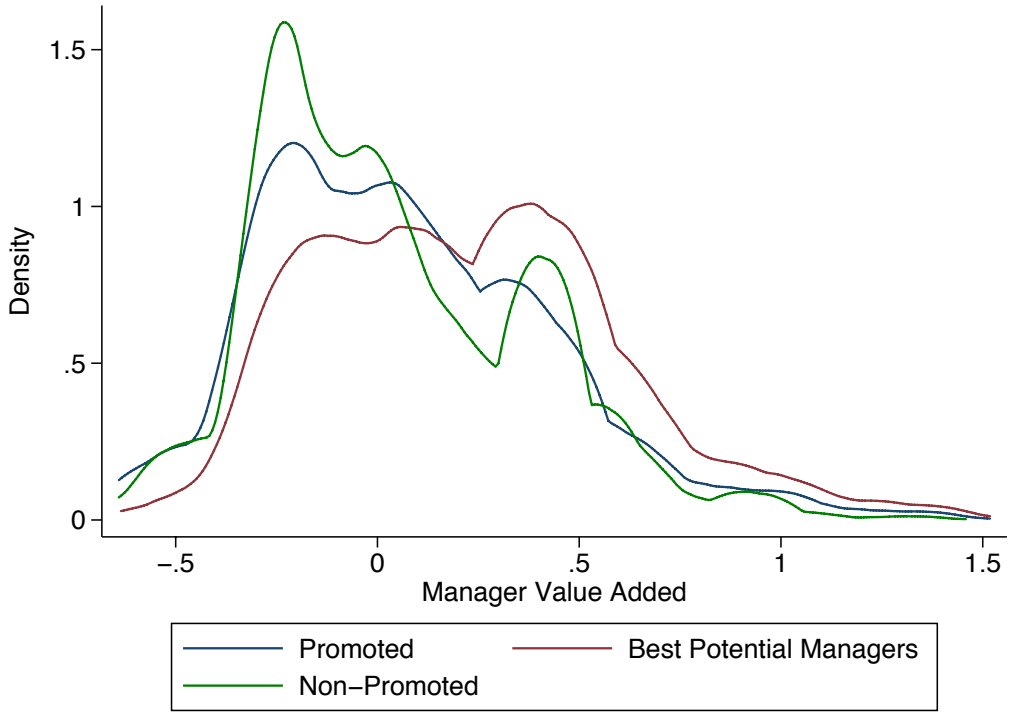
Notes: The top panel shows the kernel density of the predicted values of promotion propensity for salespeople above and below the median for residualized sales, where residualized sales is estimated as the residual from a regression of log sales on firm-month fixed effects. The bottom panel shows the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

Figure 4: Manager value added and promotion propensity



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 top panel shows this relation separately for the sample with above/below median residualized sales (after controlling for firm fixed effects). The bottom 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: 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 (optimally promoted). All manager value added measures are predicted fitted values of manager fixed effects regressed on a five part spline in log sales, log collaborators, team size, and tenure at the time of the relevant promotion event. To estimate optimal promotion decisions, 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

Note: * 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 fixed effects	654

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

Note: 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 worker promotions

	Worker is promoted	
	(1)	(2)
Log(sales)	0.0181***	0.0323***
	(0.00181)	(0.00295)
Log(collaborators)	0.0126***	0.00491
	(0.00434)	(0.00707)
Tenure and team size FEs	Yes	Yes
Company-Month FEs	No	Yes
R-squared	.001	.056
Clusters (workers)	47254	47254
Observations	698965	698965

Standard errors in parentheses

* p¡0.10, ** p¡0.05, *** p¡0.01

Note: 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 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. The other covariates include fixed effects for seven bins of a worker's tenure (and tenure interacted with a censored dummy) and team size. Team size is distinct from our main measure of collaboration experience and represents the number of other sales workers who share the same manager (team members do not necessarily collaborate on sales transactions). Column 2 also includes controls for company-month fixed effects. Standard errors are allowed to be clustered by worker.

Table 4: Correlates of Manager Value Added

	Manager value added, promotions to all positions		Manager value added, only promotions to non-parent positions	
	(1)	(2)	(3)	(4)
Log(sales)	-0.0864**	-0.0883**	-0.0977***	-0.101**
	(0.0353)	(0.0437)	(0.0364)	(0.0454)
Log(collaborators)	0.258**	0.270*	0.252**	0.256
	(0.111)	(0.142)	(0.120)	(0.160)
Company FEs	No	Yes	No	Yes
R-squared	.016	.073	.017	.076
Observations	2231	2231	2143	2143

Standard errors in parentheses

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

Note: This table presents the regression described in Equation (3). We use data at the manager level. 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 year prior to promotion (or for i's total pre-promotion tenure, if less than 12 months). Odd numbered columns include controls for company fixed effects. Columns 3 and 4 further restrict the sample to managers who are assigned to subordinates that are not their previous teammates. To more efficiently estimate the firm fixed effects, the sample includes both managers who have been promoted, for whom we observe pre-promotion characteristics, as well as those who have always been managers during our sample period. For those who have always been managers in our sample, log sales and log collaborators are missing and set to zero, and we introduce an indicator for a non-promoted manager as an additional control variable. Standard errors are adjusted for heteroskedasticity.

Table 5: Over- and Under-weighting in Promotion

	Manager value added, promotions to all positions				Manager value added, only promotions to non-parent positions	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(sales)	-0.0728**	-0.0932**	-0.0641*	-0.0903**	-0.0776**	-0.105**
	(0.0338)	(0.0366)	(0.0381)	(0.0441)	(0.0392)	(0.0461)
Log(collaborators)		0.258**		0.285**		0.294*
		(0.108)		(0.137)		(0.156)
Promotion propensity						
I	0.111	-0.0399	0.0985	0.0344	-0.0431	-0.145
	(0.310)	(0.305)	(0.401)	(0.395)	(0.432)	(0.425)
$I^2 * 10^1$	-0.252	0.0605	-0.444	-0.407	-0.0690	0.0347
	(0.611)	(0.604)	(0.844)	(0.841)	(0.916)	(0.911)
$I^3 * 10^2$	0.226	-0.00946	0.505	0.506	0.179	0.141
	(0.447)	(0.441)	(0.646)	(0.646)	(0.701)	(0.701)
$I^4 * 10^3$	-0.0761	-0.00484	-0.189	-0.193	-0.0827	-0.0774
	(0.133)	(0.131)	(0.198)	(0.198)	(0.215)	(0.215)
$I^5 * 10^4$	0.00792	0.000863	0.0208	0.0214	0.00985	0.00962
	(0.0131)	(0.0129)	(0.0199)	(0.0199)	(0.0215)	(0.0215)
Company FEs	No	No	Yes	Yes	Yes	Yes
R-squared	.015	.02	.073	.078	.076	.08
Observations	2231	2231	2231	2231	2143	2143

Standard errors in parentheses

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

Note: This table presents the regression described in Equation (10). We use data at the manager level. The variables $I_i + \cdots + I_i^5$ represent a quintic for manager i's promotion propensity as a worker. All other variables and sample restrictions are as defined for Table 4. Standard errors are adjusted for heteroskedasticity.

Table 6: Assignment of Managers to Subordinates

	Subordinates' sales, past 7-12 months		Subordinates' sales, past 4-6 months		Subordinates' sales, past 1-3 months	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(sales)	0.232***		0.257***		0.213***	
	(0.0443)		(0.0466)		(0.0492)	
Log(collaborators)	0.00116		-0.0139		0.00854	
	(0.135)		(0.147)		(0.158)	
Manager value added		0.0274		0.00239		-0.0463
		(0.0691)		(0.0960)		(0.0892)
Company FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.664	.76	.647	.747	.633	.731
Observations	1357	1357	1357	1357	1357	1357

Standard errors in parentheses

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

Note: 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 4. Manager value added fixed effects are also as defined for Table 4. Standard errors are adjusted for heteroskedasticity.

Appendix Table A: Over- and Under-weighting in Promotion: Restricted Manager Sample

	Manager value added			
	(1)	(2)	(3)	(4)
Log(sales)	-0.0678**	-0.0867**	-0.130**	-0.151**
	(0.0342)	(0.0369)	(0.0596)	(0.0665)
Log(collaborators)		0.247**		0.227
		(0.107)		(0.181)
Promotion propensity				
I	0.141	-0.00650	1.422***	1.383***
	(0.289)	(0.281)	(0.513)	(0.508)
$I^2 * 10^1$	-0.318	-0.0103	-3.589***	-3.605***
	(0.572)	(0.558)	(1.223)	(1.230)
$I^3 * 10^2$	0.276	0.0440	3.622***	3.669***
	(0.422)	(0.410)	(1.245)	(1.259)
$I^4 * 10^3$	-0.0913	-0.0210	-1.481***	-1.503***
	(0.127)	(0.123)	(0.556)	(0.562)
$I^5 * 10^4$	0.00943	0.00246	0.208**	0.210**
	(0.0125)	(0.0121)	(0.0892)	(0.0901)
Company FEs	No	No	Yes	Yes
R-squared	.021	.036	.202	.207
Observations	656	656	656	656

Standard errors in parentheses

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

Note: This table reestimates Columns 1-4 of Table 5, but restricts the sample to managers for whom we have both pre-promotion performance data as well as estimated managerial fixed effects. In other words, other managers are not included to more precisely estimate the firm fixed effects.

Appendix Table B: Over- and Under-weighting in Promotion: Outliers Winsorized

	Manager value added			
	(1)	(2)	(3)	(4)
Log(sales), winsorized	-0.0732**	-0.0923**	-0.0642**	-0.0896***
	(0.0338)	(0.0364)	(0.0389)	(0.0665)
Log(collaborators), winsorized		0.256***		0.283***
		(0.107)		(0.136)
Promotion propensity				
I	0.151	-0.00286	0.132	0.0633
	(0.289)	(0.282)	(0.379)	(0.372)
$I^2 * 10^1$	-0.338	-0.0173	-0.522	-0.472
	(0.572)	(0.559)	(0.806)	(0.802)
$I^3 * 10^2$	0.290	0.484	0.566	0.556
	(0.422)	(0.412)	(0.622)	(0.622)
$I^4 * 10^3$	-0.0913	-0.0220	-0.208	-0.208
	(0.127)	(0.123)	(0.192)	(0.192)
$I^5 * 10^4$	0.00978	0.00254	0.0226	0.02228
	(0.0126)	(0.0122)	(0.0193)	(0.094)
Company FEs	No	No	Yes	Yes
R-squared	0.015	0 .020	0.074	0.078
Observations	2,231	2,231	2,231	2,231

Standard errors in parentheses

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

Note: This table reestimates Columns 1-4 of Table 5, but uses variables winsorized at the top and bottom 2 percent levels in both the first and second stage regressions.