

C E



**Center for
Effective
Organizations**

**Optimizing Incentive Plan Design:
A Case Study**

**CEO Publication
T 11-10 (598)**

Alec Levenson

*Center for Effective Organizations
Marshall School of Business, USC*

Cindy Zoghi

Bureau of Labor Statistics

Michael Gibbs

*University of Chicago Booth School of
Business & IZA*

George Benson

*College of Business Administration
University of Texas at Arlington*

September 2011

Optimizing Incentive Plan Design: A Case Study*

Alec Levenson
Center for Effective Organizations
Marshall School of Business, USC
alevenson@marshall.usc.edu

Cindy Zoghi
Bureau of Labor Statistics
zoghi.cindy@bls.gov

Michael Gibbs
University of Chicago Booth School of Business & IZA
michael.gibbs@ChicagoBooth.edu

George Benson
College of Business Administration
University of Texas at Arlington
benson@uta.edu

September 1, 2011

We study effects of a firm's attempt to optimize an existing incentive scheme to increase sales growth for direct store delivery workers. Before optimization workers reported Ratchet Effects that lowered productivity. The altered incentive plan offered higher compensation for increased sales relative to a sales growth target, and lower compensation for failing to meet the target. We gathered data on performance and attitudes at pilot and control sites before and after the change. Relative to control sites, sales growth increased in the pilot sites by two percent, a meaningful contribution to firm profits. We find no change in distortion of effort or manipulation of the performance measure. Workers did not substantially change number of hours worked, though allocation of time across tasks changed slightly. Despite increased productivity, workers continued to report Ratchet Effects after the change. We also find that an unplanned price increase midway through a fiscal year affected the extent of Ratchet Effects that year.

JEL codes: M52; J33; M12; L81.

* All views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics. We thank participants at the 9th IZA/SOLE Transatlantic Meeting of Labor Economists, the 2010 CAED/COST Conference, and the 2011 SOLE Annual Meeting for helpful comments. Gibbs gratefully acknowledges funding from Chicago Booth and the Center for Corporate Performance, Aarhus School of Business.

I. Introduction

The role of compensation in driving behavior has been a central concern for economics and management researchers for many years. Despite this, there is scant experimental analysis of responses to changes in incentive-based compensation systems as firms try to optimize their design, and of Ratchet Effect labor supply responses to incentive systems. We provide evidence of both, using a case study of direct store delivery sales representatives.

Several studies have found productivity effects from introducing pay for performance. Seiler (1984) finds that workers paid by piece rate earn higher pay due in part to higher effort, and in part to a risk premium. Lazear (2000) finds a 44 percent productivity increase in a windshield installation company, and attributes about half of this to incentives and half to sorting. Paarsch and Shearer (1999, 2000) find a similar productivity differential among tree planters paid piece rates, relative to those paid a fixed salary. In a laboratory experiment, Cadsby, Song and Tapon (2007) find that incentive pay leads to higher productivity among all subjects, and that more risk-averse individuals prefer a fixed salary and are less productive under incentive pay.¹

The large productivity differentials found in these studies highlight the potential benefits of performance pay, which helps explain the popularity of such compensation schemes for workers whose individual output is easily observable. Design and implementation of even simple incentive schemes may be complex. First, it is difficult to set the correct initial incentive intensity, since the firm is likely to have imperfect information on how productive workers can be. It may then adjust the incentive intensity once it observes performance in response to the initial incentive. If workers anticipate this, they may reduce effort to avoid having the rate lowered in the next period. This is known as the “Ratchet Effect” (Roy 1952, Mangum 1962, Stiglitz 1975, Frexias, Guesnerie and Tirole 1985, Lazear 1986, Gibbons 1987), which may occur if the firm cannot commit ex ante to a fixed rate. The spinning industry in 19th Century England was able to avoid the Ratchet Effect through inter-firm collusion to keep piece rates fixed (Laznick 1981, Huberman 1986). The Lincoln Electric Company has been able to pay piece rates for approx-

¹ Harrison, Virick and William (1996) find higher turnover when sales representatives switch from a base pay plus commission reward structure to a commission-only pay system.

imately ninety years by developing a reputation for never lowering the piece rate, paying a subjective annual bonus, offering guaranteed employment and other supporting policies (Milgrom and Roberts 1995), but Lincoln is a highly unusual company in both policies and history.

In this study, we examine the effects of a change from a linear incentive scheme to a slightly more complex scheme, using personnel data from a large consumer products company. Owan and Tsuru (2010) examine a change in incentive pay structure similar to that observed here, for sales workers in a large Japanese auto company. We study direct store delivery workers who are initially paid a straight commission based on sales on their routes. Survey data indicate that these workers regularly withhold output in anticipation of having their routes changed to take away volume. The firm conducted a pilot study to test a change to the program designed to strengthen incentives and increase sales. The new compensation scheme offered a steeper incentive rate for sales around the target, plus a new base pay for low sales. Compared to the initial pay plan, compensation was higher if sales exceeded the target, and lower if sales did not.

Our research design has several advantages. First, the experimental evaluation design in a real world setting provides direct and credible estimates of employee responses to changes in performance pay. Second, the changes implemented are typical of how firms optimize performance pay for workers who already are paid on incentives. Third, to our knowledge this is the first recent empirical study to provide statistical evidence of Ratchet Effects at the individual worker level, and to explore the relationship of the Ratchet Effect to measures of labor supply, productivity, and changes in the compensation system.

Prior to the program launch, average sales growth in pilot sites was 1% lower than in control sites. After the launch, average sales growth in pilot sites exceeded that in control sites by 1%. Both differences are statistically and economically significant. These results may not seem dramatic compared to those of other studies. However, other studies compared implementation of incentives to straight salary. We are studying the effects of marginal adjustments to incentives designed to optimize an existing incentive plan's design. The results indicate that marginal changes to incentives lead to marginal changes in effort. Moreover, the effect of two percent greater sales growth on profit is non-trivial to the firm. Absent an extraordinarily high new product introduction rate, large consumer products companies typically ex-

pect sales to grow at or slightly above the national growth rate in consumer nondurable spending. Consumer nondurable spending growth historically ranged in the low single digits.² Thus two percentage points of additional sales growth can add substantially to both revenue and profit growth. We find that workers did not significantly change total hours worked, but slightly changed the allocation of time across main job tasks. These results indicate that one way in which incentive schemes affect behavior is through the intensive margin – working harder and more effectively for a given amount of time on the job – and not just the extensive (time) margin.

We find some of the first empirical evidence of Ratchet Effects in incentive systems since Roy’s famous paper, both before and after the compensation system change. Employees report that they sometimes hold back performance to avoid a greater increase in next year’s target. We also find that an unexpected price increase midway through a fiscal year affected the extent to which employees held back performance to avoid ratcheting up of their target in the next year.

II. Company, Compensation Plans and Data

The company studied is in a consumer products industry and uses a direct store delivery (DSD) system.³ Under DSD, manufacturing companies take responsibility for shipping products from their plants and warehouses directly to retailers, and for stocking the product on the shelves (versus using the retailer’s warehouse and distribution systems). The focus of this study is the Route Representative (which we often refer to as a “DSD employee”), who is responsible for four tasks: ordering products, delivering them to the retailer, stocking retailer shelves, and negotiating additional shelf space. With the exception of delivering the product, all of the other tasks have a direct impact on sales, and require the driver to intelligently analyze sales trends, manage relationships with store managers, and creatively market the product within the store. Ordering affects product mix and avoids out-of-stock situations, stocking (location, neatness

² Between 2000 and 2010, the annual growth rate of nondurable consumer spending ranged from -1.2% and 3.7% with an average of 2.05 (calculation from Bureau of Economic Analysis data, Table 2.3.1: Percent Change From Preceding Period in Real Personal Consumption Expenditures by Major Product Type, www.bea.gov).

³ The data are confidential, including the identity of the firm, region, or specific products it sells. We cannot report actual sales, compensation, number of employees, etc., though we have full access to all detailed information. Some figures reported here are disguised (e.g., by adding or subtracting a constant to change the mean value).

of the display, rotating older products to the front, etc.) affects how customers choose which product to buy, and gaining additional shelf space for displaying the product (at the expense of competing products) affects total volume sold.

Retail establishments sell a wide variety of products, so each manufacturer has minimal impact on the flow of traffic into a store. Instead, competition for consumer dollars at the retail level takes place among products that are substitutes in the consumer's shopping cart: tobacco competes with alcohol, various drinks compete with each other, breakfast cereal competes with ready-to-eat breakfast items, over-the-counter weight loss aids compete with comfort foods, etc. At the margin, manufacturers increase sales through branding and marketing campaigns, price promotions, gaining additional shelf space at the retail establishment, and varying the product mix available on the shelves. Of these, DSD employees can directly influence the last two.⁴

The firm groups routes in a local geographic area into "sites" with common management oversight. A site consists of a contiguous geographic area covering a distinct consumer and labor market. For example, a small- or medium- sized city and surrounding communities might constitute a site. Each site has a leadership team, supervisors, and DSD drivers. To keep the span of control and complexity at a reasonable level, each site ranges from approximately 80 to 120 routes. Very large urban areas contain multiple sites, but the area covered by each site within those large urban areas is drawn to correspond to natural community divisions.

A. Initial and New Pay Schemes

Figure 1a illustrates the original and pilot incentive schemes. Both are linear commissions on the dollar value of net sales revenue (gross sales net of returned / unsold product). There are a number of advantages to this general approach in both the original and pilot schemes. A linear scheme based on net revenue is easy to understand by employees and supervisors, and to administer. There is transparency because there is little dispute regarding the dollar volume of total sales: employees observe which prod-

⁴ Prices are set centrally rather than by DSD employees or store managers. That is true for most products sold in these types of retail stores.

ucts sell at what prices, and which do not and are returned to the warehouse. The performance measure should provide relatively well-balanced multitask incentives and thus not distort incentives much away from optimal behavior. As long as profit margins do not vary much across individual company products, revenue provides incentives to manage tradeoffs between sales of different products in the same store. It also motivates the employee to effectively allocate effort to different tasks. The measure also motivates the employee to avoid excess inventory, since unsold product is not included in sales. (The employee is not charged for the opportunity cost of excess inventory in the store or truck, but that is likely to be a small cost for this firm's products.)

The company had two concerns with the original pay scheme, which they tried to address with the pilot scheme. First, the original scheme did not explicitly motivate DSD employees to care about sales growth, on which management's incentives are based. Second, the firm wanted to increase the strength of the incentive. In addition, although not an explicit concern of the firm, the incentive plan may motivate Ratchet Effects from how targets are set. We examine all three of these issues in the empirical work. Next, we describe the initial pay scheme in detail – many of its features were maintained in the pilot scheme – and then how the firm adjusted it in the pilot scheme.

i. Initial Pay Scheme

The commission in the initial pay scheme is fixed and depends on route size. The company divides DSD employees into three route-size categories. Within each category, it attempts to design specific routes – a set of stores in the same geographical location – so that expected compensation will fall in a fairly narrow range, assuming equal ability and effort of employees. The commission rate for each route type is set to equalize expected compensation across the three types, so the highest volume routes have the lowest commission rate, and vice versa.

Though employee compensation under the initial pay scheme does not depend on a sales target, bonuses for frontline supervisors and leadership depend nonlinearly on the site meeting its growth targets. This was a main source of disconnect between the company's objectives and DSD employees' incentives under the initial compensation scheme. Sales targets are set for the DSD employees too, in an attempt to

encourage them to meet growth targets, and supervisors regularly met with the employees to identify ways to increase sales. However, under the initial pay scheme DSD employees are rewarded directly for increased sales, but only indirectly for sales growth, and their commission does not vary with the growth rate.

Setting of sales targets for employees at various levels is done using a complex combined top-down and bottom-up approach, which we now describe in some detail. This approach applies to setting target sales in both the initial and pilot compensation plans.

First, dollar volume sales targets are set for senior management, based on the total sales growth target set by the company. Those cascade down to site targets for the middle and frontline managers through the annual budget and planning process. Sales growth rate targets for supervisors are determined by targets for their DSD employees. Though differences across routes might in principle lead to different sales growth targets among supervisors, in practice individual route differences equal out across supervisors. Thus supervisors have essentially the same sales growth targets as the overall site target (i.e., if the site's target is 2 percent, the target for each supervisor is also 2 percent).

Sales targets for DSD employees are first and foremost determined by the site's overall target volume sales: if the site's target growth rate is 3 percent, individual route targets are set to yield 3 percent growth. Thus the highest volume routes have the largest dollar volume targets, and vice versa. Deviations from that distribution are allowed in unusual circumstances. For example, sales growth at new retail outlets are historically higher than at older establishments, so a route with a high concentration of older outlets might receive a slightly lower sales target. Such adjustments to a route's target are made by site leadership (not the supervisor), but only by making offsetting changes in targets for other routes at the site.⁵ This process ensures that gaming of targets by supervisors is minimized, and that adjustments to individual route targets are the exception and not the rule.⁶

⁵ Compliance with this process was ensured by the company's software used for target setting, which automatically proportionately reassigns any volume target taken away from one route to all the other routes at a site. For example, if there are N routes and X dollars are subtracted from one route's target, on average $X/(N-1)$ dollars are distributed to the targets for other routes in that area. In practice the distribution is proportionate to historical route volume so that that adjustment to the sales growth target is the same across all routes.

⁶ Compared to a system where supervisors have more direct control over the setting of targets. In systems where supervisors'

Relative performance of the employee may be accounted for indirectly because further adjustments are possible in cases where historical sales growth is significantly lower than expected given route characteristics. Employees deemed to have provided relatively low effort might be given higher targets in the subsequent year. In practice, however, the single most important determinant of an employee's target sales is the overall growth goal for that region.

Note that this approach seems to imply a Ratchet Effect: an employee's sales this year raise the target sales level next year. However, under the initial compensation scheme, the commission rate is constant, and compensation is therefore independent of whether or not the employee meets the growth target. Ratchet Effects may instead be created by the firm's practice of route "rebalancing." If expected sales on a route grow too large, the job may become difficult for one employee to handle because of the physical requirements that the employee deliver and stock the products sold. In such cases the firm adjusts the set of stores in the employee's route to bring volume back in line with the company's view of physical limitations of DSD employees in performing the job. Lower volume also makes it easier for sales growth targets to be achieved, though at the cost of creating more routes in aggregate. Such rebalancing typically takes place at most once per year. Route rebalancing is one of the most disliked and distrusted management processes. The company uses lump-sum pay adjustments to try to ease the impact on total compensation, but the adjustments are imprecise. As a result employees might exhibit a Ratchet Effect whereby they lower the growth rate of sales to try to avoid larger adjustments through route rebalancing.

Now consider the intensity of incentives in the initial plan. DSD jobs typically have high sales volume of small, inexpensive products. It is not uncommon for sales on a route to average \$10,000 per week or more, for total annual sales around \$500,000. DSD employees typically are high school graduates or have some college education. Assuming target compensation of approximately \$40,000, the commission rate would need to be 8 percent. Higher expected sales would mean an even lower commission rate. The nature of the job and high volume of products sold mean that relatively little time is spent on pure sales activities such as negotiating for additional display space on a regular basis. Thus a key question is

compensation depends on the collective performance of their direct reports and the supervisors can influence target setting, they have an incentive to set targets low to ensure that the employees can exceed them.

whether the commission rate in the initial plan provides sufficient incentive to elicit effort. For example, additional sales of \$500/week, added up over all employees' routes, could significantly help a company meet its profit objectives. Yet at 8 percent commission, and assuming a marginal tax rate of 25 percent, an extra \$500/week in sales translates into only about \$30/week extra in take-home pay. This is one reason the company changed the compensation system.

ii. Pilot Pay Scheme

The pilot pay scheme was designed to provide stronger incentives for employees to meet or exceed their targets, so that local management would be better able to meet overall growth goals set for them by higher levels of management. This was accomplished by establishing base compensation BC, or minimum salary paid regardless of volume, and a steeper commission rate paid out only for sales growth above a threshold T, as shown in Figure 1a. Interestingly, this is precisely the kind of change modeled by Oyer (2000), with the same intention of increasing the incentive intensity in a desired range of output. For sales growth above TS, compensation under the pilot scheme was higher than under the original scheme, and vice versa for sales below TS.

The company set the new compensation formula so that compensation is equal in both plans when sales equals the growth target. The base pay is intended to provide a safety net to workers in the case of natural disasters or other calamities that might cause consumers to stop purchasing products other than bare necessities for extended periods of time. In normal circumstances, workers are expected to achieve at least minimum expected sales $M > T$, so that base pay BC will be non-binding. While a worker will not face termination for performing below M in a single month, consistently performing below M will trigger termination (this was also true under the old system). In the pilot plan, the worker knows BC, T, and the commission rate, but M is implicit.

One potential concern in sales incentive plans is that employees might time sales or returns of items around the end of a bonus period (e.g., end of fiscal year) in order to maximize bonuses across the two years (Oyer 1998). In principle that concern might apply to the pilot scheme, for employees with

performance very close to Threshold Sales T , since sales below T receive no commission. However, employees with sales so low are highly likely to be terminated, eliminating that concern.

By changing the relationship between pay and sales growth, the new incentive scheme had the potential to change Ratchet Effect incentives. Under the original scheme, an employee would want a route with the highest volume possible to maximize total sales, even though growth might be minimized. The threat of route rebalancing created a potential Ratchet whereby employees might restrict sales growth to avoid having an account taken away. Under the new scheme, with compensation aligned with sales growth instead of sales volume, employees on lower volume routes should have an easier time meeting sales growth targets than employees on higher volume routes (assuming the same growth rate target). Thus under the new pay plan employees might no longer resist route rebalancing, and might even welcome it if they have confidence they can more easily grow sales with the stores remaining on a route after rebalancing. This does not ensure a lower Ratchet Effect under the new scheme, but suggests that it might be lower under the right circumstances.

B. Pilot Inception and Data Collection

The company decided to initially change the compensation scheme at two of their sites. Two of the authors, Benson and Levenson, were engaged by the company to conduct an external evaluation of the compensation change as a project of the Center for Effective Organizations at the University of Southern California. They had no role in design of the pilot incentive scheme. The external evaluation entailed working closely with corporate and local management to identify issues to be addressed by the evaluation, conduct interviews at pilot and control sites after launch of the new scheme, and design and field two surveys to measure before-and-after impacts of the change.

The sites chosen for the pilot program were representative of the middle of the company's distribution of growth performance. They were selected to be as normal as possible, meaning that sites with known management problems or peculiar demand or labor market characteristics (e.g., a highly depressed local economy) were excluded. For each pilot site, two control sites were chosen – one that matched as closely as possible in terms of pre-pilot performance, and one that did not necessarily match as well in

pre-pilot performance but had demand characteristics (socioeconomic and demographic) very similar to the pilot site.

Prior to initiating the change, a survey was distributed for the workers to fill out in private and return by mail to the external evaluators. Workers at pilot and control sites were not informed of plans to experiment with a revised pay plan, nor that they were considered to be pilot or control sites. This is important, because it means that pilot employees did not alter their behavior in anticipation of a higher commission rate. Workers were given paid time to fill out the surveys and were assured that their confidentiality would be maintained. All subsequent matching with archival data was performed by the external evaluators. Only summary information from the surveys that could not be linked back to any individual's responses was reported to the company as part of the evaluation process.

1,008 surveys were distributed and 746 were returned, for a response rate of 74 percent. Table 1 shows that of these 746 respondents, 272 were workers in the pilot sites, and 474 were in the control sites. Additionally, only 611 (216 in pilot sites and 395 in control sites) of these are considered "primary" workers. Higher volume routes are staffed by a primary person who works only that route five days per week, and by a secondary person who is responsible for the other two days per week (ensuring the higher volume routes are staffed seven days/week) and splits time between two higher volume routes. The primary person has the bulk of control over performance on the route, so we limit the analysis to primaries only. We also exclude new hires and others who are not dedicated to specific routes and spend their time filling in on routes where there is a temporary vacancy.

The survey obtained basic worker demographic information: age, race, gender, marital status, number of children, and education level; years of experience in the current position, in the company at any position, outside the company in other DSD jobs, and general experience of any kind; and information on second-job holding and commuting. Additional questions included hours worked each week, and how that time was spent between a number of job tasks. Finally, a large number of Likert-scaled questions measured workers' risk aversion, motivation, selling techniques, relationship and interactions with supervisor, organizational support, understanding of the sales goals and pay scheme, sense of pay equity, job satisfaction, pay satisfaction, job security and intention to quit. These questions were selected from estab-

lished constructs used within the management research literature to analyze these issues (e.g., Park and Holloway 2003, Fang, Evans and Zou 2004; Tekleab, Bartol and Liu 2005).

The compensation change was introduced to the pilot sites in July and August 2007, after all the surveys had been returned. Management held a site-wide meeting to announce the new pay system, and supervisors subsequently spent many hours answering workers' questions. After giving the new pay scheme almost two years to affect workers' performance and attitudes, the survey was repeated in May 2009. This round, 986 surveys were distributed, and 724 of them were returned, for a response rate of 75 percent. As Table 1 indicates, 231 of the second round respondents were in pilot sites, and 493 were in control sites. Among the 586 that were considered primary route workers, 187 were in pilot sites, and 399 were in control sites. As the bottom rows of Table 1 show, of those who returned the round one survey (746), 608 received a round two survey, with the remainder having left the company or taken a different job within the company. 486 of those returned a round two survey—149 in pilot sites and 260 in control sites. Since we are mainly interested in the difference-in-difference between round one and round two at the pilot sites compared to round one and round two at control sites, this is the primary sample for most of our remaining tables.

These survey data were then linked to personnel data by the external evaluators. Sales volume and growth was provided by the company for the specific retail establishments on the worker's route, relative to the same establishments in previous periods. The average percent of plan attained was provided, from which the number of times the workers achieved their sales plans was calculated. Additionally, supervisor ratings of the workers' sales performance, as well as driving and delivery performance and merchandising/customer service performance, all from the supervisor surveys, were matched to the survey data.

Between administration of the first and second surveys, in the middle of 2008, the company instituted a price increase that was unusual in both size and timing. Price increases are usually introduced around the beginning of the calendar year, which facilitates setting sales targets at the corporate level that are incorporated at all levels from senior leadership down to DSD employees. Input price pressures had been building for a while, but the company had previously resisted fully passing them along to customers

for fear of losing market share. By mid-2008, the recession and slower sales threatened attainment of corporate profit goals. The company decided to put through a relatively large and extensive price increase that covered a much broader set of products than the usual end-of-year adjustments. Target sales (TS) for employees were not changed. The increase had not been accounted for in route planning for the year, and represented a surprise to local management and DSD employees. Questions about the impact of the price increase on employees' ability to meet sales targets were included in the second round survey. At face value it appears that the price increase should have affected employees in pilot and control sites equally.

C. Descriptive Statistics

Table 2 provides summary statistics. We see that the demographics of pilot and control employees are very similar; the only statistically significant difference is that pilot workers are slightly younger. DSD workers are predominantly male. Around 35 percent have a high school education, 46 percent have some college, and more than 16 percent have a Bachelor's degree or more. Around 8-10 percent are non-white. The employees have fairly high job tenure,⁷ averaging 11 years with the company in round one, increasing to 12½ to 13 years by round two. This is partly due to the exclusion of new hires, who are not primaries on a route. This is also a job that encourages high tenure to some extent, because seniority determines priority on bidding for route assignments. Primary route employees are around 42-43 years old in round one, and have 24-25 years of general work experience, 5 of which is as a DSD worker for another company. While Table 2 shows that these numbers change significantly between round one and round two as expected, any differences between pilot and control groups are insignificant.

Table 3 shows that the workers in the pilot sites seemed to understand clearly how the new pay scheme affected their incentives. When asked how important growth potential and high volume are to their choice of routes, they initially respond similarly to those in the control sites. After the pilot introduction, however, they indicated a much stronger preference for growth potential and a lower preference for high volume, consistent with the incentives of the new pay scheme. Workers in pilot sites significantly

⁷ According to the BLS (<http://www.bls.gov/news.release/tenure.nr0.htm>), median job tenure for all U.S. workers was 4.0 years in 2006 and 4.1 years in 2008.

increased their stated commitment to their sales goals, which was an intention of the new incentive program. Workers in the pilot sites also slightly increased their tolerance for risk in terms of route preference. Despite the added complexity of the new pay structure, there was no significant difference in the pilot workers' understanding of their goal before the pilot start versus after.

Workers in the pilot sites were less likely than those in control sites to agree that goals were fair after introduction of the pilot plan, despite answering similarly in round one. We take this as evidence of the heightened importance of the goals in the employees' minds, since there was no major change in the goal-setting process. If anything, the fairness of the process might have increased because management knew that they had to be even surer to get the process right once the compensation scheme changed. They consequently increased their efforts in the pilot sites post-implementation to ensure that the site-wide sales growth goal was distributed across routes in accordance with underlying business conditions and growth potential on each route. (This was confirmed through interviews with corporate and local management at each site.) Thus the increase in stated unfairness of the goals most likely reflects their heightened importance under the new compensation scheme (see also the discussion of goal importance consequences below).

III. Empirical Predictions

We now briefly discuss likely effects of the pilot program compared to the original incentive scheme to frame the empirical analysis. Figure 1b focuses on the portion of Figure 1a in the range where virtually all employees will operate, since those with performance too close to M are likely to be terminated quickly. Figure 1b shows indifference curves for an employee whose initial performance is exactly at target sales TS . The indifference curves point toward the northwest since lower output implies less effort/ higher leisure. At that point, the employee's optimum effort and output will increase by switching to the pilot pay plan, precisely the effect the firm intended. Moreover, the employee's utility will rise. If an employee's output were initially to the right of that point, at output higher than TS , utility would also rise with the pilot program. The opposite would hold if output were initially below TS . In both of those latter cases the effect on output is ambiguous because it is possible that income effects dominate substitution

effects. However, that seems unlikely since the change in total compensation in going from the initial pay plan to the pilot is small – the substitution effect is likely to dominate the income effect – and the firm intended the change to increase sales. Therefore, we expect higher performance in the pilot plan than in the initial plan.

DSD jobs involve multitasking as described above. Therefore, increased total performance could come via the extensive margin of greater hours worked, or changes in allocation of time to different tasks. If observed, such changes would not be driven by changes in relative weights on various tasks in the incentive scheme because the incentive scheme rewards outputs, not inputs. However, it is possible that relative marginal disutilities of effort, or marginal products of effort, change with total worker output. We examine both total output and employee multitasking in the empirical work.

Lazear (1996, 2000) found that the range of output increased when Safelite implemented pay for performance. We will examine whether similar effects occur in the pilot study. In Lazear’s case the logic was that low ability workers will be more likely to quit, and high ability workers more likely to apply for a job, after Safelite implemented piece rates. Our logic is not based on selection via recruiting and turnover, but instead on the different effects of the incentive change on the *same set* of workers. In the pilot study, workers have strong incentives to achieve threshold sales each year. If they can achieve that level, they then have marginal incentives to increase sales incrementally. Thus, performance might become more skewed to higher output. This would be especially likely if effort and ability are complements in production.

Lazear also found significant turnover from Safelite’s new pay scheme. We expect turnover to be highest when making large changes, such as from an hourly pay rate with no incentive component to a piece rate system, as in Safelite. In the present case, we expect the “fit” story to be less important because compensation both before and after the change was based on incentive pay tied to the volume of sales. What changed was the relative strength of incentives. There are two reasons for potentially increased turnover as a result of the compensation scheme change. Lower productivity workers who cannot meet the new performance standards, and stand to receive a cut in compensation, might choose to leave voluntarily. Involuntary turnover might also increase if the company couples the new compensation scheme

with an increased performance management emphasis on meeting the growth targets, treating below-average performance as a possible cause for termination. We test for these effects by examining the productivity of leavers in the pilot versus control sites, and the workers' and managers' responses to questions about termination risk. We also check whether leavers had lower performance and growth before leaving.

Ratchet Effects are difficult to model formally, especially with closed-form solutions, even for linear incentive schemes (Gibbons 1987).⁸ To illustrate why this is the case, consider how this firm sets the commission b in the second period as a function of performance in the first period, $b_{t+1} = TC/TS_{t+1}$. Ratchet Effects will occur if TS is a function of prior output. Suppose a simple model for output, $sales_t = a \cdot e_t + \varepsilon_t$, where a = marginal product of effort or employee ability, e = effort, and ε = performance measurement error. Further assume, close to what the firm actually does, that TS is a simple percentage increase r over last period output, $TS_{t+1} = (a \cdot e_t + \varepsilon_t)(1+r)$. If so, the second period commission is $b_{t+1} = TC/(a \cdot e_t + \varepsilon_t)(1+r)$. From the perspective of period t , b_{t+1} is a random variable, but analysis of its properties is difficult since it has the stochastic term ε_t in its denominator. However, we can see where the effect comes from right away, since $\partial b_{t+1}/\partial e_t = -aTC/[(a \cdot e_t + \varepsilon_t)^2(1+r)] < 0$. The harder the employee works in period t , the smaller the commission he will be paid in period $t+1$.

Ratchet Effects may also depend on some other variables. The greater the expected growth in target compensation TC next year, the smaller the Ratchet Effect. We have no information on worker expectations about growth in compensation during the period studied, so cannot study that prediction. If we interpret a as worker ability, then Ratchet Effects may vary with ability. We will attempt to construct proxies for variation in worker talent to examine this. Finally, the larger the firm's growth rate in target r , the weaker the Ratchet Effect.

⁸ Lazear (1986) analyses a closed-form Ratchet Effect model, but in his model the ratchet is done to base pay, not the commission rate, which makes the analysis tractable.

IV. Worker Responses to the Pilot Compensation Plan

A. Changes in Labor Supply and Effort

Given the small nature of the income effects, we expect sales to grow faster in the pilot sites than the control sites after the introduction of the new compensation scheme. Table 4A shows the median sales growth and percent of sales plan attained by workers before and after the pilot start date, at both pilot and control sites. Considerable care was taken to ensure that sales growth numbers are measured as a change in sales at the set of retail establishments on the worker's route, not as a change in the worker's sales. Otherwise, a worker on a re-engineered route might have higher sales by virtue of having a set of higher-volume retail establishments, rather than by increasing sales at the same establishments.

Sales in the pilot sites were initially growing 1.2% slower before the pilot launch, compared to the control sites. This most likely was due to idiosyncrasies of the local consumer demand in the pilot sites, because the sales growth targets essentially do not vary at the site level (in percentage terms). Indirect evidence for this can be found in the supervisor responses to the question "Employee assigned sales plan goals are very difficult to achieve" in the pre-pilot survey round: 45.5 percent of supervisors "agreed" or "strongly agreed" with this statement in the control sites, versus 61.9 percent in the pilot sites. This suggests that local demand conditions in the pilot sites pre-launch might have been slightly below their longer-run potential.⁹

During the first year of the pilot launch, sales in pilot sites grew 0.7% faster, though the gap narrowed to 0.3% faster in the second year. Taken at face value, this appears to be evidence of a positive impact of the change in incentives on effort and output.¹⁰ The second panel of Table 4A shows similar patterns for the percentage of employees who met their performance goal. Prior to the start of the pilot

⁹ Suggestive evidence in favor of this can be seen in Appendix Table A.1 which breaks out separately the sales performance for the two pilot sites. One site had significantly lower performance in the pre-period, whereas the other did not. The site with the lower performance had a more rapid increase in sales after the pilot launch, which is consistent with a return to more normal conditions. Performance in the second site nonetheless also improved, and went from slightly lower (and not statistically significantly different) than the control sites, to higher than the control sites (and statistically significant). Thus while idiosyncratic factors might account for part of the increase in sales in one of the pilot sites, the overall patterns across the two pilot sites suggest a real impact of the compensation scheme change.

¹⁰ For confidentiality reasons we cannot provide many details on site level performance. However, a two percentage point relative increase in sales for a site is very meaningful in business terms for the company.

compensation plan, 0.9% fewer pilot site employees met their goal than in the prior year, while 0.9% more control site employees met their goal. After implementation of the change in incentives, the situation reversed, with pilot sites showing significant increase in the percentage of employees that met their goal, compared to control sites. As described above, in the middle of 2008 an unexpected price change was introduced. In the period following the price change, sales growth was not significantly different at pilot sites relative to control sites. In addition, control sites showed an increase in ability to meet their goals compared to pilot sites. This suggests that the price increase made it easier for control sites to improve their performance.

Table 4B presents regression analyses similar to the summary statistics in Table 4A. Sales growth is regressed on indicators for whether the worker is in a site with the new pilot compensation scheme and time period (column 1) and worker characteristics (column 2). Thus, the coefficient on the pilot site dummy measures the difference in differences between performance before and after the pilot start, in pilot sites versus control sites. Sales growth increased after the pilot introduction one half percent faster for those in pilot sites than for those in the control sites. Worker characteristics (marital status, race, gender, and measures of human capital) had no effect on the worker's response to the change in incentives.

Figures 2a and 2b depict the distribution of average sales growth for the control group and the pilot group, respectively. The numbers represent the six months before the pilot, and the year or so after the pilot, but before the price change is introduced. To complement the patterns in Figures 2a and 2b, Table 5 reports the percentage of people in the pilot and control sites falling into the tails of the distribution (below zero percent growth and above ten percent growth, respectively) for the pre-pilot period, for the first year following the pilot launch, and for the second year after that.

Given the results of Table 4B, it is not surprising that both Figures 2a and 2b and Table 5 show that the distribution of sales growth in the pilot sites shifted to the right in the year following the pilot launch; the disappearance of the far left tail represents routes that had a decrease in sales from the past year. There is also a decrease in the density of the far right tail. However, to reach any conclusions about the overall change in dispersion at the pilot sites, the patterns need to be compared to the performance distribution changes in the control sites, which were similar. Differencing the estimates from Table 5 indi-

cates that the density of the lower tail of the performance distribution in the pilot sites decreased by 5.1 percentage points more than the control sites, while the density of the upper tail of the performance distribution in the pilot sites decreased by 4.1 percentage points less than the control sites. Thus the density in both tails of the performance distribution combined decreased by 1.0 percentage points for the pilot sites relative to the control sites. This is counter to our expectation that the range of output would be wider for the pilot workers than the control ones. However, that discussion ignored the possibility of a compensation scheme-induced change in Ratchet Effect that might have caused the top part of the performance distribution in the pilot sites to contract more than the lower parts of the distribution, compared to the control sites.

Employees can increase their labor supply either by increasing the total number of hours worked or by increasing their effort and working harder/smarter. Workers in this company have a high degree of independence, and largely set their own schedule to best accommodate the retailers they service. There are no set hours required by the company, although routes are engineered such that it would be difficult to make the deliveries working very short hours. Table 6 shows that workers averaged over 50 hours per week at both the pilot and control sites. The first column shows the allocation of those hours to different tasks. The second column estimates the effect on hours worked in each task as a function of whether the worker was in a pilot site and was in the second time period (there are two observations per worker, one for each round of the survey). These coefficients measure the difference-in-differences before and after introduction of the new compensation scheme, in pilot v. control sites. The table shows that pilot workers did not increase their hours significantly more than workers in control sites. Despite this, pilot workers increased output as described above.

Thus, we now consider whether the employees achieved faster sales growth by redistributing their effort from the less sales-oriented tasks to the more sales-oriented tasks. The surveys asked workers what percent of their time was spent among a variety of tasks. The first column of Table 6 shows that the employees spend the greatest amount of their time on merchandising and display (33-34 percent), followed by driving and delivery (around 20 percent). While workers in the pilot sites spent slightly less time on sales activities and planning pre-pilot compared to the control sites, the differences are not statistically

significant. Moreover, while both pilot and control workers increased the total time spent on planning and sales activities in round two, the increase was larger and statistically significant for the pilot sites (3.5 versus 2.4 more hours). These changes are indicative of a change in the distribution of hours worked toward more sales-related activities.

Taken together, the evidence on increased sales, total hours worked and distribution of hours worked across tasks implies that higher productivity was accomplished through workers supplying greater effort in the tasks they performed and spending more time on value-added activities — they did not work longer hours, they just worked harder and/or smarter.

B. Performance Management and Turnover

In the analyses to this point we have focused on the incentive effects of the compensation system change. In order to compare “apples to apples” and focus on the incentive effects, we have restricted the focus to people who were employed continuously as leads on the routes. Yet to obtain a complete picture of the compensation system change on overall performance at the sites, we have to account for turnover and potential productivity differences between departing and new employees at each site. We start by considering the performance management standards and any changes that might have accompanied the new compensation scheme and then turn to potential productivity impacts created by the substitution of new employees for those who left.

Because the employees have essentially no influence over the traffic of customers who come into a store, a significant portion of the sales on an account are beyond their control. The company is well aware of this and has performance management processes that ensure a minimum performance standard. For all intents and purposes this means that sales on a route cannot consistently fall below a particular level (measured in relative terms to prior year sales) without triggering termination. There is no explicitly stated threshold for triggering termination because local market factors (e.g., population growth; income growth) vary substantially, and because rare events like natural disasters can cause the sales on a route to shrink abnormally for reasons beyond the employee’s control for extended periods of time.

Though terminations happen, they are relatively rare and predominantly concentrated among low tenure employees who are still in the process of learning about the job. Terminations “for cause” among longer tenure employees are essentially indistinguishable from voluntary terminations. Once someone has worked on a route long enough, they know how to maintain performance above the termination threshold. Longer tenured employees who end up getting fired are likely intending to quit anyway and pull back on effort to enjoy more leisure time while they count the days to when they are pushed out.

Interviews with management representatives indicated that the performance management standards remained the same for the pilot sites when the new compensation system was introduced. Thus the threat of termination remained unchanged. In Table 7, we report evidence that workers in the pilot sites were more concerned about meeting their goals after the new pay scheme was introduced. The variable Goal Importance Consequences indexes several survey questions on the negative repercussions of failing to meet sales goals, including general negative consequences, career limitations, and potential termination. See Appendix Table B.1 for a complete list of the index variables constructed from the survey items. Between round one and round two, workers in pilot sites became significantly more concerned about the consequences of not meeting their goals, while workers in control sites became less so, undoubtedly due to the 2008 price change that made it easier for them to attain their sales goal.

Examining the specific question about potential termination, in the pilot sites the mean response increased by a statistically significant amount from 2.81 to 3.08 while it remained essentially unchanged in the control sites, 2.82 versus 2.86. In a separate survey administered to supervisors, the comparable question about termination risk due to missing sales plan goals was not significantly different at pilot and control sites and did not change from round one to round two, confirming the management interview results about termination risks. This suggests that the institution of a compensation penalty for sales below the target (for the new compensation scheme compared to the old one) sent a signal to the employees that the company cared about the sales growth targets. The employees in turn may have misinterpreted the compensation emphasis change to include tougher performance management standards that could lead to termination. Thus part of the behavior change in the pilot sites could have been driven not just by responses to the compensation changes but also by a perceived termination threat. Despite this perceived job

insecurity, Table 7 also shows that workers in the pilot sites were no less satisfied with their jobs or pay before the pilot introduction than after, although the price increases seem to have improved the sense of well-being of workers in the control group.

Another way that overall performance might increase at pilot sites relative to control sites is if the introduction of the pilot compensation induces turnover. Table 8 estimates the probability of leaving the company, as a function of being in one of the pilot sites. The first column indicates no higher likelihood of leaving for those workers in pilot sites, with no additional controls. Such a result might still arise, however, if the pilot were launched in sites where there was a lower propensity to leave before the pilot was introduced, and this offset any increase in turnover that would have resulted from the pilot introduction. To address this, the survey asked a number of questions about job satisfaction, including specific questions about workers' intention to turnover. A composite variable was created out of two highly correlated Likert scaled questions: "I am seriously thinking about quitting my job" and "I intend to quit my job," coded from 1 = strongly disagree, through 5 = strongly agree.

The second column of Table 8 shows how the probability of leaving is affected by intending to leave beforehand. Although the coefficient is positive, it is not significant, which either implies that workers are not fully forthright in answering such questions, or that their attitudes subsequently changed. The third column shows that controlling for the propensity to leave before the pilot is introduced does not change the result that workers in the pilot sites are no more likely to leave afterwards than those in control sites. In the fourth column, we interact intention to turnover with the pilot indicator. The coefficient on intention to turnover indicates that those in control sites who intend to leave are, in fact, more likely to do so. Those in pilot sites are on average more likely to leave, but those with a higher intention to quit are actually less likely to subsequently do so.

Columns 5 and 6 divide the sample into pilot sites and control sites to show the same relationship between intention to turnover and leaving as in column 4. Only in the case of the control sites does intention to turnover in period one lead to leaving in period two; in pilot sites, those who plan to quit are subsequently no more likely to do so after the pilot is introduced. The results for the control sites are consistent with extensive evidence in the management literature (Hom and Griffeth, 1995; Griffeth, Hom,

Gaertner, 2000) showing a positive relationship between intention to turnover and actual turnover. Thus the results for the pilot sites suggest that the introduction of the new compensation scheme may have led workers who otherwise would have quit to delay leaving, at least temporarily. The opportunity to perform under different incentives may have been sufficient to induce some of these workers to change their minds and work at the company at least a little while longer.

In assessing the productivity effects of the new compensation scheme, it is important to account for actual turnover. If workers with lower performance leave the company following the introduction of the pilot program, it might appear that the pilot enhances the productivity of workers, when in fact the remaining workers are no more productive than in the control sites. To assess this, we compare the productivity of leavers versus stayers at pilot versus control sites. Table 9 estimates ordinary least squares regressions of sales growth as a function of whether the worker left in a later period. Column 1 measures sales growth in the period before the introduction of the pilot, while column 2 measures sales growth in the period after the introduction of the pilot. Leavers are defined as those who do not have sales data for 2009. Leavers at all sites are significantly less productive in the pre-pilot period, but no more or less productive than stayers following the pilot introduction. In columns 3-6, we disaggregate the analysis into workers at pilot sites and control sites. In the control sites, leavers had lower productivity in the first period, while in the pilot sites, leavers were equally productive before the pilot introduction but less productive in the second period, following the introduction of the pilot. Columns 7 and 8 show that there are no significant productivity difference in either period between those leavers in pilot and control sites.

In addition to workers leaving the company, there are new hires in the company. The new compensation scheme may attract a more motivated worker, since it rewards high effort workers more. Thus, there may be productivity differences in new hires. Table 10 compares the post-pilot productivity for workers who did not have sales data from the pilot start through the end of 2007 and those who were present from 2007 through 2009. Column 1 estimates sales growth in the second period for all sites as a function of whether the worker was a new hire. Among all sites, joiners were no more or less productive than incumbents. Columns 2 and 3 disaggregate the analysis into pilot site workers and controls. The coefficients on joiners have opposite signs, with joiners in the control sites being a statistically signifi-

cantly 1.2% more productive; and, while not statistically significant, joiners in the pilot sites are less productive. Column 4 directly measures the difference in productivity of joiners in pilot versus control sites, by restricting the sample to joiners and including a control for whether the worker joined a pilot site. Workers who joined pilot sites were over 2% less productive than those who joined control sites.

C. The Ratchet Effect

Under the original linear commission pay scheme, the incentive for a Ratchet Effect was straightforward. With compensation based entirely on the dollar volume of sales, the route representative's incentive was to have the highest volume of sales possible without triggering a route rebalancing (which takes volume away from the route and thus directly lowers compensation). In this section we provide some evidence that employees did report engaging in Ratchet Effect behavior under the original plan, and then some brief analysis of how the incentive plan change affected such ratcheting behavior.

Table 11 presents results from an earlier survey we conducted in this company in 2005. Though the sample of employees and DSD routes is different than the rest of the tables that we focus on in this paper, we present these summary statistics because the much large sample size allows for more precise measurement of Ratchet Effects and their impact on sales under the original compensation plan. Approximately one-fifth indicated that they managed sales growth on their route to avoid rebalancing, while about one-sixth indicated they managed sales growth to avoid having the number of calls (number of times per week they had to show up at the premises of a customer)¹¹ per account increased. To better understand these self-reported Ratchet Effects and their impact on performance, we also report correlations of these measures with total sales on the route, supervisor ratings of job performance (sales, delivery/driving, merchandising), and self-reports of the employees' time spent on the different tasks, separately for high volume and for the low volume routes.

¹¹ The number of calls per account is mandated by management and used as a performance metric for the employees. Under the initial compensation scheme there was no direct impact on compensation from failing to meet that performance target, though failure to do so could be viewed as potentially slacking off and thus evidence for reduced effort on the route (which in turn could impact subsequent route rebalancing).

For both the high and low volume routes, employees who report ratcheting behavior to avoid having the number of calls per account increased (columns four and eight) are rated lower on all three facets of job performance by their supervisors (with five of the six relationships being statistically significant at the 90 percent confidence level or higher). There also appears to be a shifting of time spent on the job away from sales activities and toward delivery/driving activities on both route types, among those reporting a Ratchet Effect with respect to the number of calls on an account. The net result is a negative impact on sales for high volume routes, though not for low volume routes. Among employees reporting a Ratchet Effect to avoid having accounts taken away by route engineering (columns three and seven), there are no significant correlations with the supervisor rating, time allocation and sales measures for the high volume route employees. Low volume route employees reporting such Ratchet Effects are rated lower by their supervisors on the performance of all three job facets.

Taken together, the patterns in Table 11 indicate that self-reported Ratchet Effect measures can be used to identify behaviors consistent with a Ratchet Effect, including lower sales. They also provide insights into the challenge of the original compensation plan, and the firm's motivation for implementing the pilot compensation system change to increase the employees' incentives to provide marginal effort and labor supply. Under the new compensation scheme, the effect of rebalancing on an employee's ability to maximize compensation changed from being a clear negative to being indeterminate. While an employee's target sales fall with rebalancing, the percentage growth in that target over the prior year might be larger or smaller with rebalancing, depending on which stores are kept in or eliminated from the route. Thus under the pilot the employees might report decreased Ratchet Effect compared to the original compensation scheme.

Table 12 reports survey responses to two sets of questions about Ratchet Effects for employees in the pilot and control groups studied in the rest of the paper other than Table 11. The top panel shows that there is a fairly substantial Ratchet Effect before the compensation scheme change, and that it is higher at pilot sites: almost a quarter (23.5 percent) of employees at control sites indicate that either they or someone they know manage the sales growth on their routes to avoid having accounts taken away by route rebalancing, compared to more than a third (35.7 percent) of employees at the pilot sites. The responses in

round two suggest a (statistically insignificant) drop in Ratchet Effects at the pilot sites for self-reported ratchet (A) from 20.6% to 16.3% while the control sites had a smaller decline from 9.6% to 8.1%. This is consistent with a potentially greater willingness of pilot employees post-launch to undergo route rebalancing, though the statistical weakness of the results means this conclusion is only speculative. The differences in initial self-reported Ratchet Effects could be due to differences in trust of the route rebalancing process at the two groups of sites. Consistent with that, initial pay satisfaction and perceived pay fairness is lower at the pilot sites than at the control sites (Table 7). A separate set of questions on the round two survey measured Ratchet Effects in response to the unexpected 2008 mid-year price increase. Similar to the pattern in the top of the table, the pilot sites' employees are more likely to report "slacking off" because the price increase made it easier to meet the sales plan goals.

Finally, Table 13 checks for differences in sales growth between employees, based on whether or not they reported Ratchet Effects. There is no evidence in the top panel that the changes in the incentive system led those who reported Ratchet Effects to have significantly different sales growth than those reporting no Ratchet Effect. This is not surprising, as it was unclear from a theoretical perspective whether the pilot plan should unequivocally have any such effect. Table 13 does reveal, though, that the unexpected price increase in mid-2008 led to lower sales growth for those who had previously reported that they engaged in ratchet behavior or were aware of colleagues who did. The price increase made it easier to meet the sales goal in 2008, and may have been perceived as increasing the likelihood that an employee would face route rebalancing at the beginning of 2009. This may have induced marginal Ratchet Effects to try to avoid rebalancing.

V. Conclusion

In this study, we examine the effects of a change from a linear piece rate scheme to a more complex compensation scheme, using personnel data from a large consumer products company. We also provide evidence of Ratchet Effects and their correlates at the individual level. We study direct store delivery workers who initially were paid a straight commission based on sales on their routes. In a number of locations, a new payment scheme was introduced which offered higher incentives for increased sales relative

to the previous year, as well as higher penalties for failing to meet sales objectives. Workers at both treatment and control sites were surveyed both before and after the change was introduced, gathering information on attitudes and self-reported Ratchet Effects.

Prior to the program launch, average sales growth in the pilot sites was one percentage point lower than in the control sites. After the pilot launch, average sales growth in the pilot sites exceeded that in the control sites by one percentage point. Both differences are statistically and economically significant. The workers did not substantially increase hours worked, though they slightly changed the distribution of time spent on different job tasks. The evidence from changing the compensation scheme suggests that the original linear incentive pay did not elicit maximum worker effort, and that the new compensation scheme elicited greater effort.

References

- Cadsby, C. Bram, Song, Fei and Tapon, Francis (2007). "Sorting and incentive effects of pay for performance: An experimental investigation," *Academy of Management Journal*, 50:2, 387-405.
- Fang, Eric, Evans, Kenneth R., and Zou, Shaoming (2004). "The Moderating Effect of Goal-Setting Characteristics on the Sales Control Systems – Job Performance Relationship," *Journal of Business Research*, 58, 1214-1222.
- Frexias, Xavier, Roger Guesnerie and Jean Tirole (1985). "Planning Under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52, 173-191.
- Gibbons, Robert (1987). "Piece-Rate Incentive Schemes," *Journal of Labor Economics*, 5:4(1), 413-429.
- Griffeth, Roger W., Peter W. Hom, and Stefan Gaertner (2000). "A Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator Tests, and Research Implications for the Next Millenium," *Journal of Management*, 26:3, 463-488.
- Harrison, David A., Virick, Meghna, and William, Sonja (1996). "Working without a net: Time, performance, and turnover under maximally contingent rewards," *Journal of Applied Psychology*, 81:4, 331-345.
- Hom, Peter W. and Roger W. Griffeth (1995). *Employee Turnover*, Cincinnati, OH: South-Western.
- Huberman, Michael (1986). "Invisible Handshakes in Lancashire: Cotton Spinning in the First Half of the Nineteenth Century," *Journal of Economic History*, 46:4, 987-998.
- Lazear, Edward P. (1986). "Salaries and Piece Rates," *Journal of Business*, 59:3, 405-431.
- Lazear, Edward P. (1996). "Performance Pay and Productivity," National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5672, July.
- Lazear, Edward P. (2000). "Performance Pay and Productivity," *American Economic Review*, 90:5, 1346-1361.
- Lazonick, William (1981). "Competition, Specialization and Industrial Design," *Journal of Economic History*, 41:1, 31-38.
- Mangum, Garth (1962). "Are Wage Incentives Becoming Obsolete?," *Industrial Relations*, 2, 73-96.
- Milgrom, Paul and John Roberts (1995). "Complementarities and Fit: Strategies, Structure, and Organizational Change in Manufacturing," *Journal of Accounting and Economics*, 19, 179-208.
- Owan, Hideo and Tsuyoshi Tsuru (2010). "Integrating High-Powered Performance Pay into a Seniority Wage System." Working paper, University of Tokyo.
- Oyer, Paul (1998). "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *Quarterly Journal of Economics*, 113(1): 149-185.
- Oyer, Paul (2000). "A Theory of Sales Quotas with Limited Liability and Rent Sharing," *Journal of Labor Economics*, 18(3), 405-426.
- Paarsch, Harry J. and Shearer, Bruce S. (1999). "The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree Planting Industry," *Journal of Human Resources*, 34:4, 643-667.
- Paarsch, Harry J. and Shearer, Bruce S. (2000). "Fixed Wages, Piece Rates, and Incentive Effects: Statistical Evidence from Payroll Records," *International Economic Review*, 41:1, 59-92.
- Park, Jeong-Eun and Holloway, Betsy B. (2003). "Adaptive Selling Behavior Revisited: An Empirical Examination of Learning Orientation, Sales Performance, and Job Satisfaction," *Journal of Personal Selling & Sales Management*, 23:3, 239-251.
- Roy, Donald (1952). "Quota Restriction and Goldbricking in a Machine Shop," *American Journal of Sociology*, 57:5, 427-442.

- Seiler, Eric (1984). "Piece Rate vs. Time Rate: The Effect of Incentives on Earnings," *Review of Economics & Statistics*, 66(3), 363-376.
- Stiglitz, Joseph E. (1975). "Incentives, Risk and Information: Notes Toward a Theory of Hierarchy," *Bell Journal of Economics*, 6:2, 552-579.
- Tekleab, Amanuel G., Bartol, Kathryn M. and Liu, Wei (2005). "Is It Pay Levels or Pay Raises that Matter to Fairness and Turnover?" *Journal of Organizational Behavior*, 26, 899-921.

Figure 1A. Pilot and Control Pay Schemes

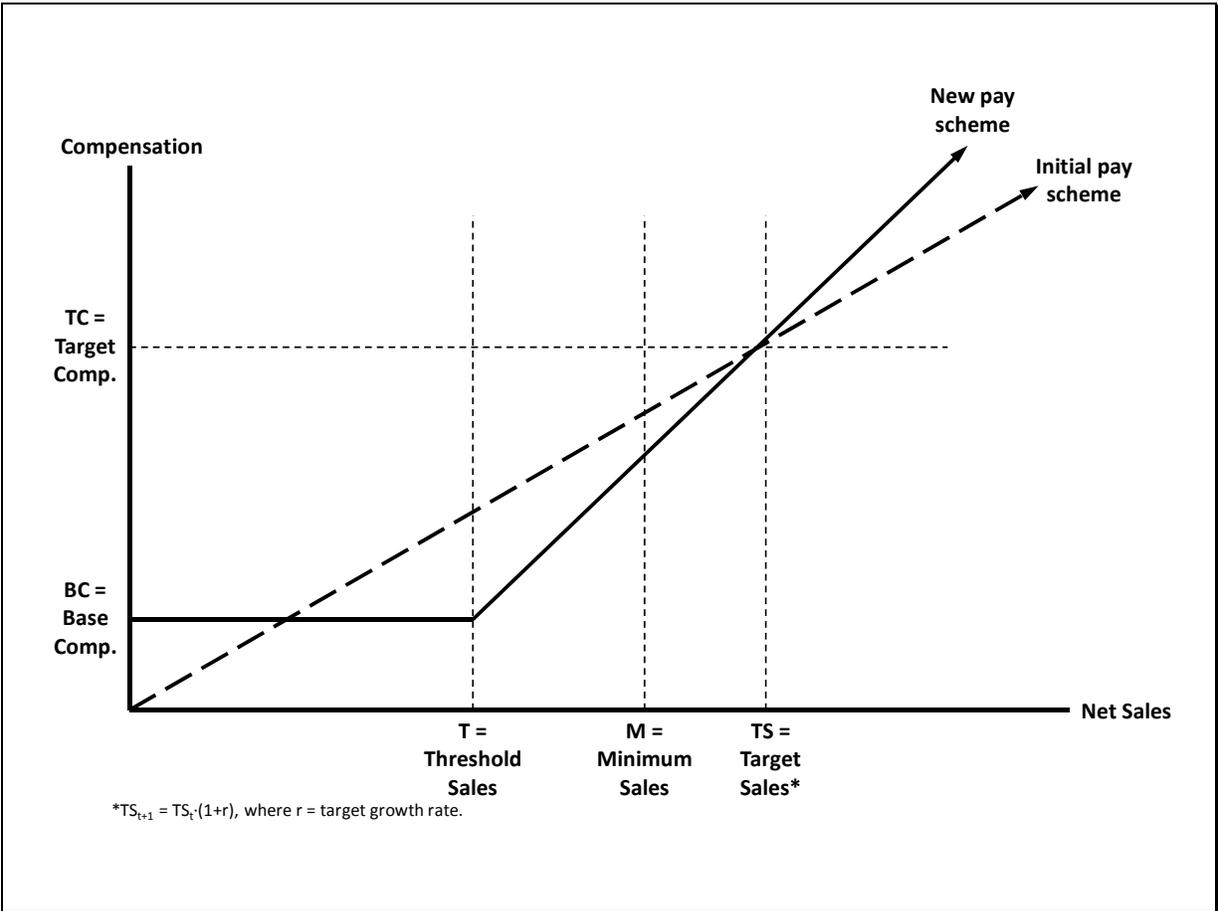


Figure 1B. Incentive Effects of Pilot Scheme

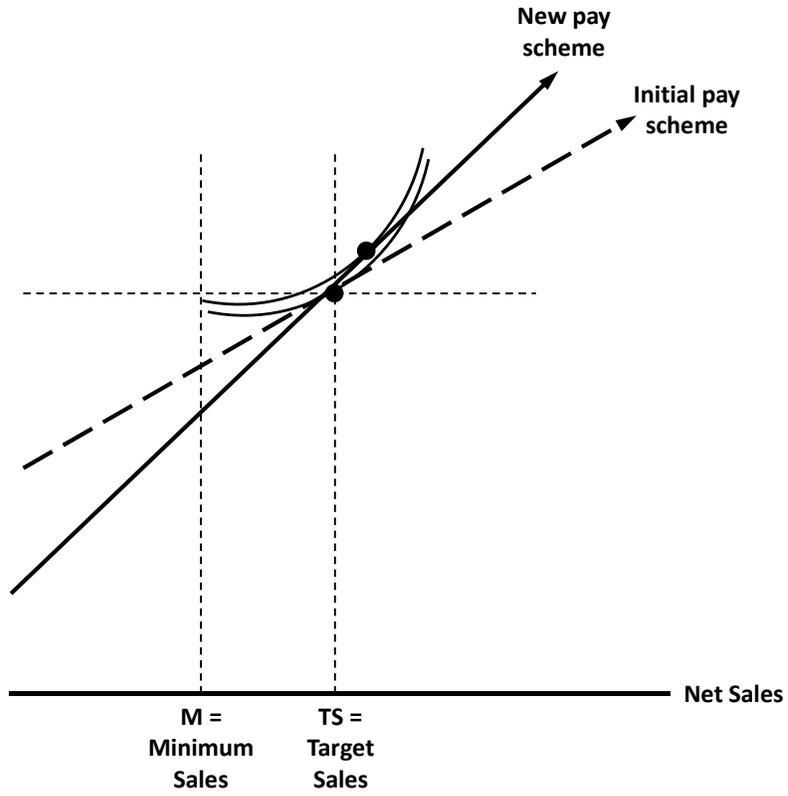


Figure 2A.
Average sales growth before & after pilot start, control group

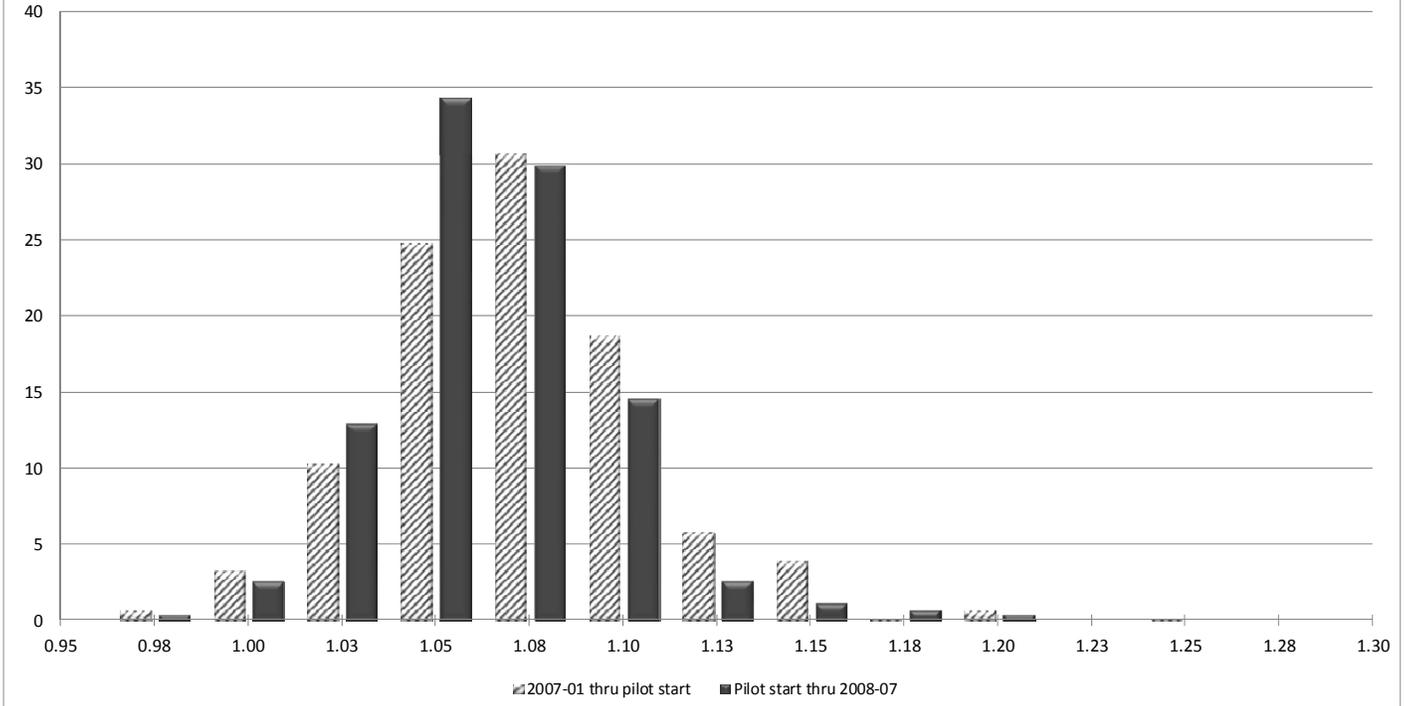


Figure 2B.
Average sales growth before & after pilot start, pilot group

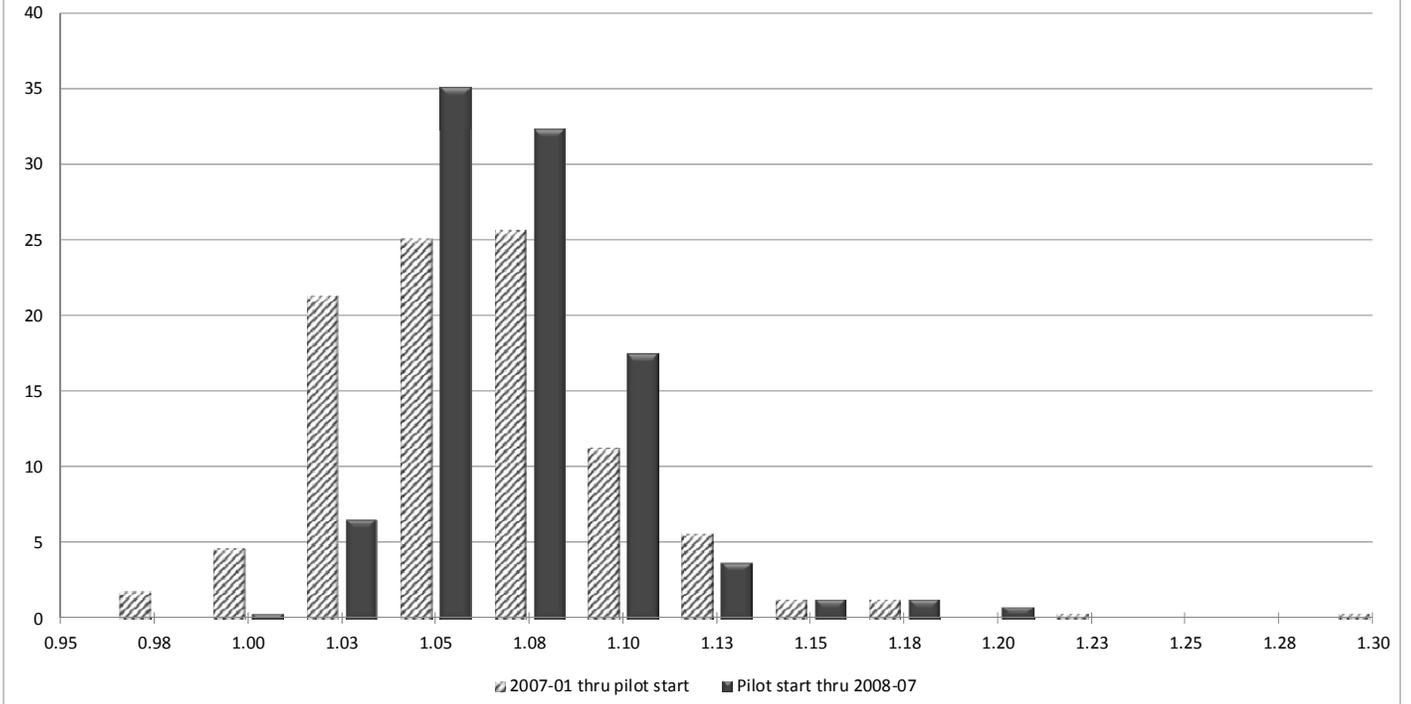


Table 1. Sample sizes			
	All Employees	Pilot	Control
Full sample			
All workers surveyed in either round	1,194	375	819
Round 1 (April 2007) survey period			
All workers surveyed	1,008	318	690
Respondents	746	272	474
Primaries ^a on route	611	216	395
Round 2 (July/August 2008) survey period			
All workers surveyed	986	326	660
Respondents	724	231	493
Primaries ^a on route	586	187	399
Both rounds			
All workers surveyed in both rounds	800	269	531
Respondents in both rounds	486	183	303
Primaries ^a on route both rounds	409	149	260
^a Higher volume routes are staffed by a primary person who works only that route five days a week, and by a secondary person who splits time between two higher volume routes. For most specifications, the sample is restricted to survey respondents who were primaries on a route in both survey rounds.			

Table 2. Characteristics of DSD Workers

	Round 1		Round 2	
	Pilot (N=149)	Control (N=260)	Pilot (N=149)	Control (N=260)
Job tenure (years)	11.0 ^{✓✓✓}	10.8 ^{✓✓✓}	13.1	12.6
Work experience since HS (years)	23.8 ^{✓✓✓}	25.2 ^{✓✓✓}	26.0	27.5
DSD experience outside company (years)	4.6	5.5 ^{✓✓}	5.2	6.4
Age	42.0 ^{*✓✓✓✓}	43.8 ^{✓✓✓✓}	43.8 ^{**}	45.8
High school only	36.9	35.0	37.6	35.4
Some college education	45.8%	47.0%	45.5%	46.4%
Four-year college education or greater	16.0%	17.0%	15.9%	16.5%
Non-white	8.5%	7.5%	11.3%	9.8%
Female	4.2%	4.7%	4.9%	6.5%

Includes only employees who were primaries in both rounds, and who responded in both rounds.

Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01

Significant difference between Round 1 and Round 2: ✓p ≤ .10 ✓✓p ≤ .05 ✓✓✓p ≤ .01

Table 3. Workers' opinions of the compensation plans				
	Round 1		Round 2	
	Pilot (N=149)	Control (N=260)	Pilot (N=149)	Control (N=260)
Using the numbers 1 (most important) through 6 (least important), please rank order the importance of the following route characteristics in determining your ideal route:				
Growth Potential	2.21 ^{✓✓✓}	2.16	1.80 ^{**}	2.09
High Volume	2.15 ^{✓✓✓}	2.15	3.80 ^{***}	2.14
Response scale: 1='Strongly Disagree', 2='Disagree', 3='Neither', 4='Agree', 5='Strongly Agree'				
Goal Fairness ^a	2.85 ^{✓✓✓}	2.82 ^{✓✓✓}	2.37 ^{***}	3.08
Goal Commitment	3.88 ^{✓✓✓}	3.97	4.07	3.99
I understand exact level of my assigned performance goals	3.73	3.72 [✓]	3.74	3.82
I prefer a low risk/high security route with steady sales over a route that offers high risks	3.08 [✓]	3.21	3.24	3.12
I prefer to remain on a route that has problems rather than take risk of working a new route	2.56 [✓]	2.60	2.73	2.67
Includes only employees who were primaries in both rounds, and who responded in both rounds.				
Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01				
Significant difference between Round 1 and Round 2: [✓] p ≤ .10 ^{✓✓} p ≤ .05 ^{✓✓✓} p ≤ .01				
Note: Significant difference (p ≤ .01) between Growth Potential rank (1.80) and High Volume rank (3.80) for Round 2 Pilot sites.				
^a Goal Fairness is an index, obtained through factor analysis, of 3 similarly-worded survey questions about perceived fairness of goal setting procedures. Goal Commitment is a similar index based on 4 questions about a worker's commitment to achieving the sales goal.				

Table 4A. Median Sales Growth and Plan Attainment		
	Pilot (N=210)	Control (N=426)
Sales Growth (compared to same periods prior year)		
Sales growth from beginning of 2007 through pilot Start (mid-2007)	4.7%***	5.9%
Sales growth from pilot start (mid-2007) through mid-2008	5.6%***	4.9%
Sales growth from mid-2008 through mid-2009	5.8%	5.5%
Plan Attainment		
% Plan attainment from beginning of 2007 through Pilot Start (mid-2007)	-0.9%	0.9%
% Plan attainment from pilot start (mid-2007) through mid-2008	3.0%***	1.4%
% Plan attainment from mid-2008 through mid-2009	3.8%***	3.3%
Includes only employees who were primaries in both rounds.		
Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01		

Table 4B. Regression Estimates of Sales Growth at Pilot vs. Control Sites			
	Sales Growth (Sample restricted to Primaries in both rounds)		
	Model 1	Model 2	Model 3
Pilot Site after implementation of new compensation system (Yes=1)	.005*** (.002)	.005** (.002)	.008*** (.002)
Pre-Pilot Time Period (Yes=1)	.004** (.002)	.005** (.002)	
End of Pilot Time Period (Yes=1)	.002 (.002)	.002 (.002)	
Married (Yes=1)		.004* (.002)	
Non-white (Yes=1)		-.002 (.003)	
Female (Yes=1)		.004 (.004)	
Some College (Yes=1)		-.001 (.002)	
Bachelor's Degree or higher (Yes=1)		-.004 (.003)	
Total years of experience working at Company		-.000 (.000)	
Years of work experience since high school		-.000 (.000)	
Years of prior route experience prior to joining Company		-.000 (.000)	
(Constant)	1.052*** (.001)	1.055*** (.004)	1.054*** (.001)
Adjusted R ²	.003	.005	
N	1905	1645	1905
Fixed Effects Model	No	No	Yes

*Standard errors in parentheses. Significance: ***p<=.01, **p<=.05, *p<=.10*

Table 5 Percentage of Low and High Sales Growth Performers		
	Pilot (N=210)	Control (N=426)
Low Sales Growth (below zero percent growth) compared to same periods prior year		
from beginning of 2007 through Pilot Start (mid-2007)	6.7	4.2
from Pilot start (mid-2007) through mid-2008	0.5***	3.1
from mid-2008 through mid-2009	2.4	3.1
High Sales Growth (above ten percent growth) compared to same periods prior year		
from beginning of 2007 through Pilot Start (mid-2007)	9.5	11.0
from Pilot start (mid-2007) through mid-2008	7.6	5.0
from mid-2008 through mid-2009	6.2	6.4
Includes only employees who were primaries in both rounds.		
Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01		

Table 6. Fixed Effects Estimates of Hours Worked and Time Spent on Various Tasks at Pilot vs. Control Sites		
	Mean	Coefficient on Pilot Site After Implementation (Indicator Variable)
Average hours worked per week	51.03	0.759 (.565)
Considering the total hours you work per week, what percentage of your time is spent on each of the following (total must add up to 100%):		
Planning	9.33	0.619 (1.074)
Delivery/Driving (<i>excluding sales activities</i>)	20.25	0.704 (2.082)
Sales (<i>excluding delivery/driving activities</i>)	15.19	1.514 (1.636)
Merchandising/Customer Service	34.60	1.255 (2.452)
Meetings, training, coaching	5.57	0.665 (.622)
Synching data and completing paperwork	5.80	-0.700 (.565)
Working product in your remote location (<i>if applicable</i>)	6.32	-1.085 (.839)
Shopping distribution center for product replenishments (<i>if applicable</i>)	0.85	-0.360 (.359)
Other	2.09	-2.612* (1.395)
Includes only employees who were primaries in both rounds, and who responded in both rounds. <i>Standard errors in parentheses. Significance: ***p<=.01, **p<=.05, *p<=.10</i>		

Table 7. Worker's Attitudes – Goal Importance Consequences and Job Aspects				
	Round 1		Round 2	
	Pilot (N=149)	Control (N=260)	Pilot (N=149)	Control (N=260)
Response scale unless otherwise noted: 1='Strongly Disagree', 2='Disagree', 3='Neither', 4='Agree', 5='Strongly Agree'				
Goal Importance Consequences	3.00 ^{✓✓✓}	3.10 ^{✓✓}	3.34 ^{***}	3.00
Intention to Turnover	2.34	2.21 ^{✓✓✓}	2.20 ^{***}	1.89
Pay Satisfaction	2.45 ^{*✓}	2.63 ^{✓✓✓}	2.68 ^{***}	3.14
Pay Equity [‡]	2.78 ^{***}	3.06 ^{✓✓✓}	2.94 ^{***}	3.54
Job Satisfaction	3.67	3.72 ^{✓✓✓}	3.67 ^{***}	3.94
Includes only employees who were primaries in both rounds, and who responded in both rounds. Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01 Significant difference between Round 1 and Round 2: [✓] p ≤ .10 ^{✓✓} p ≤ .05 ^{✓✓✓} p ≤ .01 [‡] 1='Very Unfair', 2='Unfair', 3='Neither', 4='Fair', 5='Very Fair'				

Table 8. Probability of leaving

Coefficient (Std. Error)	Logit Analysis of Leavers (Sample restricted to Primaries)					
	Model 1	Model 2	Model 3	Model 4	Model 5 Control Sites	Model 6 Pilot Sites
Pilot Sites	-.279 (.216)		-.143 (.271)	1.390** (.671)		
Intention to Turnover		.164 (.119)	.165 (.119)	.363** (.144)	.363** (.144)	-.300 (.238)
Interaction: Pilot Sites & Intention to Turnover				-.663** (.278)		
(Constant)	-1.588*** (.113)	-2.395*** (.314)	-2.349*** (.324)	-2.838*** (.398)	-2.838*** (.398)	-1.448*** (.540)
Pseudo R ²	.002	.004	.005	.019	.022	.012
N	814	594	594	594	383	211
<i>Significance: ***p<=.01, **p<=.05, *p<=.10</i> <i>Leavers=Missing archival data for 2009.</i>						

Table 9. Productivity of leavers vs. stayers								
Coefficient (Std. Error)	Productivity (Sales Growth) - Leavers vs. Stayers (Sample restricted to Primaries)						Productivity (Sales Growth) of Leavers - Pilot vs. Control Sites (Sample restricted to Primaries and Leavers)	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
			Control Sites		Pilot Sites			
	Pre-pilot	Post-pilot	Pre-pilot	Post-pilot	Pre-pilot	Post-pilot	Pre-pilot	Post-pilot
Leavers (Employees with missing archival data for 2009)	-.019*** (.005)	-.007 (.004)	-.023*** (.005)	-.002 (.005)	-.009 (.008)	-.019** (.008)		
Pilot Sites							.004 (.015)	-.008 (.015)
(Constant)	1.056*** (.002)	1.055*** (.001)	1.059*** (.002)	1.052*** (.002)	1.050*** (.003)	1.061*** (.002)	1.037*** (.008)	1.050*** (.008)
Adjusted R ²	.019	.002	.029	-.002	.001	.020	-.008	-.010
N	806	757	555	515	251	242	121	73
<i>Significance: ***p<=.01, **p<=.05, *p<=.10</i> <i>Pilot start: mid-2007</i>								

Table 10. Productivity of joiners vs. incumbents				
Coefficient (Std. Error)	Post-Pilot Productivity (Sales Growth) - Joiners vs. Incumbents (Sample restricted to Primaries)			Productivity (Sales Growth) of Joiners - Pilot vs. Control Sites (Sample restricted to Primaries and Joiners)
	Model 1	Model 2	Model 3	Model 4
		Control Sites	Pilot Sites	
Joiners (Employees with missing archival data from pilot start (mid- 2007) thru end of 2007)	.003 (.005)	.012* (.006)	-.011 (.008)	
Pilot Sites				-.021* (.012)
(Constant)	1.058*** (.001)	1.057*** (.001)	1.059*** (.002)	1.069*** (.007)
Adjusted R ²	-.001	.005	.004	.050
N	763	522	241	42
<i>Significance: ***p<=.01, **p<=.05, *p<=.10</i>				

Table 11. Correlation Evidence of Self-Reported Ratchet Effects from Earlier Survey in 2005

Correlation Significance	High Volume - Primary (N=584)				Low Volume - Primary (N=901)			
	Mean	Std. Dev.	You manage account sales growth to avoid having <u>accounts</u> taken away by route engineering	You manage account sales growth to avoid having the <u>number of calls</u> per account increased	Mean	Std. Dev.	You manage account sales growth to avoid having <u>accounts</u> taken away by route engineering	You manage account sales growth to avoid having the <u>number of calls</u> per account increased
You manage account sales growth to avoid having <u>accounts</u> taken away by route engineering	.18	.39			.19	.39		
You manage account sales growth to avoid having the <u>number of calls</u> per account increased	.13	.34	.511*** .000		.15	.36	.623*** .000	
Total Route Sales	---	---	-.035 .393	-.091** .027	---	---	-.011 .751	.001 .981
Supervisor Rating: Sales performance [‡]	3.67	1.00	-.018 .669	-.077* .063	3.61	.89	-.107*** .001	-.096*** .004
Supervisor Rating: Delivery / Driving performance [‡]	3.85	.85	-.030 .468	-.018 .671	3.85	.81	-.095*** .004	-.091*** .006
Supervisor Rating: Merchandising / Customer Service performance [‡]	3.67	1.03	-.031 .462	-.108*** .009	3.66	.94	-.059* .075	-.055* .100
Pct of time spent on Sales (excluding Delivery/Driving activities)	52.4	26.3	.013 .760	-.083** .046	48.0	24.5	-.017 .605	-.063* .060
Pct time spent on Delivery/Driving (excluding Sales activities)	41.2	25.7	-.021 .620	.052 .209	47.1	24.9	-.007 .843	.060* .072
Percentage of time spent on other activities	6.5	13.2	.015 .722	.063 .126	4.86	10.5	.056* .095	.004 .903
Significance Level: * $p \leq .10$ ** $p \leq .05$ *** $p \leq .01$								
[‡] Response Scale: 1=Worst or bottom 20%, 2=Below average or next 20%, 3=Average or next 20%, 4=Above average or next 20%, 5=Best or top 20%								

Table 12. Self-Reported Ratchet Effects Before and After the Compensation Change				
% indicating "yes"	Round 1		Round 2	
General Ratchet Effect	Pilot (N=149)	Control (N=260)	Pilot (N=149)	Control (N=260)
(A) Do you ever manage your sales growth to avoid having accounts taken away by route rebalancing?	20.6%***	9.6%	16.3%**	8.1%
(B) Do you know of other employees who manage their sales growth to avoid having accounts taken away by route rebalancing?	27.8%	20.8%	29.05**	18.0%
Either self (A) or "know of others" (B)	35.7%**	23.5%	33.3%**	21.0%
Ratchet Effect in response to 2008 price increase			Pilot (N=104)	Control (N=164)
(C) Due to the 2008 price increases, I slacked off because it was easier to make plan			6.7%	3.0%
(D) Due to the 2008 price increases, I know of other employees who slacked off because it was easier to make plan			15.5%**	7.3%
Either self (C) or "know of others" (D)			19.4%**	9.8%
Includes only employees who were primaries in both rounds, and who responded in both rounds (top and bottom of table). Sample for bottom of table further restricted to employees who responded to both sets of questions (general Ratchet Effect + Ratchet Effect in response to 2008 price increase).				
Significant difference between Pilot and Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01				
Significant difference between Round 1 and Round 2: ✓ p ≤ .10 ✓✓ p ≤ .05 ✓✓✓ p ≤ .01				

Table 13. Difference in Sales Growth between Those Who Do and Do Not Report Ratchet Effects, Before and After the Compensation Change

	Round 1		Round 2	
General Ratchet Effect	Pilot (N=149)	Control (N=260)	Pilot (N=149)	Control (N=260)
(A) Difference in sales growth between those who do and do not report ever managing their sales growth to avoid having accounts taken away by route rebalancing?	.0023	.0018	-.0002	-.0085
(B) Difference in sales growth between those who do and do not report knowing of other employees who manage their sales growth to avoid having accounts taken away by route rebalancing?	-.0059	-.0046	.0089	-.0052
Difference in sales growth between those who do and do not report either self (A) or “know of others” (B)	-.0031	.00003	.0029	-.0056
Ratchet Effect in response to 2008 price increase			Pilot	Control
(C) Difference in sales growth between those who do and do not report slacking off due to the 2008 price increases, because it was easier to make plan			.0030	.0006
(D) Difference in sales growth between those who do and do not report knowing of other employees who slacked off due to the 2008 price increases, because it was easier to make plan			-.0091*	-.0098**
Difference in sales growth between those who do and do not report either self (C) or “know of others” (D)			-.0074	-.0082**
Includes only employees who were primaries in both rounds, and who responded in both rounds (top and bottom of table).				
Significant difference between those who do and do not report Ratchet Effects: *p ≤ .10 **p ≤ .05 ***p ≤ .01				

Appendix Table A.1 Median Sales Growth and Plan Attainment				
	Pilot			Control
	Overall (N=210)	Site 1 (N=106)	Site 2 (N=104)	Overall (N=426)
Sales Growth (compared to same periods prior year)				
Sales growth from beginning of 2007 through pilot Start (mid-2007)	4.7%***	3.8%	5.5%✓✓✓	5.9%†††
Sales growth from pilot start (mid-2007) through mid-2008	5.6%**	5.4%	5.6%	4.9%†, ‡
Sales growth from mid-2008 through mid-2009	5.8%	5.3%	6.2%✓✓	5.5% ‡‡
Plan Attainment				
% Plan attainment from beginning of 2007 through Pilot Start (mid-2007)	-0.9%**	-3.8%	3.3%✓✓✓	0.9%†††
% Plan attainment from pilot start (mid-2007) through mid-2008	3.0%***	2.2%	3.5%✓✓	1.4%†††, ‡‡‡
% Plan attainment from mid-2008 through mid-2009	3.8%***	3.7%	3.9%	3.3%†, ‡‡
Includes only employees who were primaries in both rounds.				
Kolmogorov-Smirnov Tests:				
Significant difference between Overall Pilot and Overall Control sites: *p ≤ .10 **p ≤ .05 ***p ≤ .01				
Significant difference between Pilot Site 1 and Pilot Site 2&7: ✓p ≤ .10 ✓✓p ≤ .05 ✓✓✓p ≤ .01				
Significant difference between Pilot Site 1 and Overall Control sites: †p ≤ .10 ††p ≤ .05 †††p ≤ .01				
Significant difference between Pilot Site 2 & 7 and Overall Control sites: ‡p ≤ .10 ‡‡p ≤ .05 ‡‡‡p ≤ .01				

Appendix Table B.1 Survey Index Variables and Items
GOAL FAIRNESS
The steps involved in determining my sales goal are fair
The procedures used to set my sales plan goal are fair
The overall process for setting my sales plan goals is fair
GOAL COMMITMENT
I am strongly committed to pursuing the sales plan goal
It is very important to meet my sales plan goals
*Quite frankly, I don't care if I achieve the sales plan goal or not
I am willing to put forth a great deal of effort beyond what I'd normally do to achieve the sales plan goal
GOAL IMPORTANCE CONSEQUENCES
I could lose my job if I fail to meet my sales plan goals regularly
There are negative consequences if I fail to meet my sales plan goals
Failing to make my sales plan goals could limit my career at Company
INTENTION TO TURNOVER
I am seriously thinking about quitting my job
I intend to quit my job
I often think about quitting my job at Company
PAY SATISFACTION
I am satisfied with the size of my current salary
I am satisfied with my take-home pay
I am satisfied with my overall level of pay
PAY EQUITY <i>1='Very Unfair', 2='Unfair', 3='Neither', 4='Fair', 5='Very Fair'</i>
Size of your pay
Your current pay
Overall level of pay
JOB SATISFACTION
In general, I like working here
*In general, I don't like my job
All in all, I am satisfied with my job
<i>Response scale unless otherwise noted: 1='Strongly Disagree', 2='Disagree', 3='Neither', 4='Agree', 5='Strongly Agree'</i>
<i>* Item reversed for scale computation</i>