Is Cash King for Sales Compensation Plans? Evidence from a Large Scale Field Intervention

Abstract

The pervasive use of merchandise (i.e., non-cash) incentives in sales compensation plans is an empirical and theoretical puzzle given the supposed superiority of cash incentives in the standard theory (i.e., principal-agent models) and the scant, and contradictory empirical evidence. We conducted a large scale field intervention that switched 580 salespeople at a large frozen food manufacturer away from their cash plus " merchandise points" bonus to a commensurate all-cash bonus. After controlling for salesperson, seasonality, year, and target effects, we estimated that sales, on average, dropped by 4.36%. Further, we estimated individual-level sales changes and effort changes to validate our incentive-effort-sales causal chain. Our results show that the top salespeople experienced the largest drops. A post-intervention survey of social and individual difference variables reveals that salespeople from households with more discretionary financial resources, and those who think more abstractly about the uses of cash income exhibited smaller reductions in effort and sales. While the absence of a control group prevents us from making strong causal inferences, this set of results nevertheless provides descriptive and suggestive evidence for separate mental accounts as the most promising explanation for the greater utility provided by merchandise incentives.

Keywords: incentives, non-monetary compensation, field experiments, salesforce, mental accounting

Introduction

Salesforce compensation constitutes an enormous category of expenditures, totaling around $800 billion annually in the US recently (e.g., Steenburgh and Ahearne, 2012). Of this total, a significant amount is spent on incentive compensation. Incentive awards include three different reward types; cash, merchandise and symbolic awards. While both cash and merchandise rewards are financially costly to the firm, symbolic awards (e.g., recognition plaques) cost virtually nothing, so the published expenditure totals are driven by the cash and merchandise elements. While cash constitutes the majority of incentive compensation costs, one recent estimate puts the annual expenditures on merchandise incentive programs at around $80 billion (Incentive Federation, 2013) covering about three quarters of all firms. Merchandise programs grant points to salespeople for achieving sales targets, with the accrued points generally redeemable for merchandise[1] from catalogs.

Given the pervasiveness and economic importance of merchandise incentive programs, it is surprising how little we know about them. In fact, the very existence of merchandise rewards programs is something of a puzzle for standard theory, because at first glance, cash incentive programs should always constitute a lower cost solution to the firm over any combination of cash plus merchandise incentives. This conclusion flows from a consideration of the motives of the salesperson as well as those of the firm. Observe that a salesperson knows her own consumption utility better than her employer, and thus is able to make better choices for herself given a cash payout than the inherently more limited choices for redeeming her accrued points from a merchandise catalog.[2] The higher utility of the cash payout induces more salesperson effort. The firm too should strictly favor cash payouts, given the administrative costs and vendor margins associated with a merchandise incentive program. What then, from a salesperson utility perspective, is the rationale behind the existence of such programs? Drilling down, what is the relative effectiveness of cash versus merchandise incentives in eliciting salesperson effort?

Answering these questions is at the heart of this paper, and we do so through a large-scale field intervention involving salespeople of a division of a Midwestern frozen foods manufacturer.

Before describing the intervention, it is useful to examine what prior literature has to say on the rationale and effectiveness of merchandise inventive programs (and, more generally, on different types of rewards).[3]

Prior Literature

Examining salesperson preferences. The bulk of the extant work consists of stated preference studies asking salespeople about their preferences for different types of incentives (e.g., Churchill, Ford and Walker, 1985; Shaffer and Arkes, 2009; Kube et al, 2012). Regardless of the methodology (direct elicitation, conjoint), subjects consistently claim they prefer cash rewards to non-cash rewards. This stated preference for cash is exhibited uniformly across demographic categories of age and experience.

A couple of studies have elicited preferences indirectly. Shaffer and Arkes (2009) assigned laboratory subjects randomly to either a cash incentive or a merchandise incentive condition, and asked them to perform an anagram-solving task, for which they were paid (cash or merchandise) proportional to their performance. No significant performance differences were found across the two conditions. The second study is an unpublished report from a consulting firm. Gravalos and Jack (2007) describe a six-month experiment where 60 Goodyear Tire districts were randomly assigned to either a merchandise incentive program or to a cash incentive program. The programs were implemented for one line of tires carried in these districts' stores. The merchandise incentives yielded higher sales than the cash incentives. Unfortunately, it is difficult to assess this study's results as the sparse report excludes important details (e.g., cell means, variances, group equivalence, differential attrition, etc.) normally found in academic publications.

Given the conflicting evidence about the relative efficacy of cash versus merchandise incentives from these studies, we turn to several adjacent literature streams for possible insights.

Gift Wages. Akerlof (1982) denoted workplace wages above the market-clearing level as a "gift" from the employer to the employee. His model shows that a high-wage high-effort equilibrium will be sustained because workers reciprocate gifts with additional effort beyond their self-

interest maximizing level. Much like prospect theory, the key is that workers apply a reference level to ascertain their effort choices. Kube et al. (2012) recruited subjects for a study involving entering data into a library catalog. In addition to the show-up fee promised in the recruitment materials, subjects were granted an additional incentive upon reporting for the study. One group received a thermos bottle, while another group received additional cash compensation (of equal value). The authors found that the thermos gift group did more work than the cash gift group. Their explanation is that the subjects a) expend effort that they believe is appropriate, given the show-up fee, and b) separately, they expend additional effort that they believe is appropriate given the merchandise gift. The cash gift is likewise evaluated separately from the show-up fee, but it evokes less additional effort because it is summed into the show-up fee.

Extrapolating this result to our context requires caution because an element of surprise is needed to induce the subjects to perceive the additional incentive as a gift. In contrast, there are no surprises to the compensation plan for our salespeople, which is always announced ahead of the fiscal year. Likewise, tedium may set in with gifts given on a regular basis and the productivity increase from reciprocity may well fade. Indeed, Gneezy and List (2006) report that these gift reciprocity effects faded after about 3 hours on a 6-hour task using a similar task to the Kube at al. setup (but with cash gifts). In sum, these gifting studies provide no clear takeaway for our issues, but a cautious extrapolation is that merchandise incentive plans might work when the gifts are unique, and unexpected.

Non-cash Employee Benefits. Some employers (famously, Silicon Valley firms) offer non-cash benefits including beverages, meals, child care, product discounts and the like. A large literature in labor economics (e.g., Lazear, 1986; Marino and Zabojnik, 2008; Oyer and Schaefer, 2005; Oyer, 2008) finds employees value such benefits, albeit with considerable heterogeneity across demographic categories. However, there is an important distinction. These benefits are available regardless of an employee's performance, while our merchandise incentives are paid out to individual salespeople contingent on their measured sales output. As such, it is unlikely these employee benefits preference studies generalize to our setting. However, this work does foreshadow likely

preference heterogeneity across salespeople.

Loyalty Programs. Loyalty programs are ubiquitous in marketing. In a recent review, Dorotic et al. (2012) summarized the effects of a number of common features of loyalty programs, including reward tiers (volume, customer, etc.), reward networks (focal firm only, firm plus partners, etc.), and reward types. Of these, the most relevant factor for our purposes is customers' responses to different types of rewards within a loyalty program.

One strand of the literature categorizes loyalty program reward types as direct (e.g., a free coffee after nine coffee purchases) or indirect (e.g., a free haircut after five visits to a restaurant). Relating this to compensation schemes, a scheme with a base cash payment plus an additional cash award as incentive is direct; base cash plus additional merchandise awards as incentive is indirect. Direct reward types are more effective than indirect reward types (e.g., Keh and Lee, 2006, Kivetz, 2003, 2005; Yi and Jeon, 2003) which, translated to our setting, implies cash incentives are superior to merchandise incentives at inducing sales effort.

Another categorization scheme categorizes loyalty program rewards as cash-like (e.g., discounts, rebates) or non-monetary (e.g., preferred guest status) which also maps readily into our cash and merchandise incentives. Cash-like rewards are typically more effective (particularly direct cash-like rewards), but non-monetary rewards are also effective in some circumstances (e.g., Smith and Sparks, 2009; Melancon et al. 2011).

Overall, we view these loyalty program studies implying an edge for cash over merchandise incentives, but we caution that this extrapolation is subject to a large caveat. Incentive compensation aims to change effort, whereas loyalty programs seek to change consumption, which are completely different domains of action. Perhaps the most important takeaway is that cash and merchandise are perceived and evaluated in qualitatively different ways instead of being summed up first according to some exchange rate. This takeaway leads us directly to the next relevant research stream, that of mental accounting.

Separate Mental Accounts. While the loyalty program results point to the likely superiority of cash over merchandise incentives, the opposite case can be made from prospect theory (Kah-

neman and Tversky, 1979) and mental accounting (Thaler, 1985; 1999). These theories are well-established ways of describing context-dependent preference and choice structures, and they contrast sharply with the global preference structures found in traditional economics models. Briefly, mental accounting posits that our decisions are driven by a set of mental accounts that we use to segregate our wealth. For instance, one might place retirement money into one mental account and money for current consumption into another mental account, which then affects a large decision like buying a car versus increasing retirement savings.

Separate mental accounts change our evaluation of alternatives in profound ways. Evaluating two elements separately and then summing the individual evaluations yields greater utility than first summing up the two elements and then evaluating the combination. Consider applying this result to a compensation plan consisting a base element (cash), and an additional merchandise incentive. Evaluating the base and additional element separately prior to summing would lead to higher overall utility, according to the mental accounting principles described above.

The relevance of mental accounting to our setting hinges on knowing when cash and merchandise incentives might be initially evaluated separately versus initially summed up. This is addressed by Nunes and Park (2003), who argue that multiple elements of rewards and/or costs are more likely to be separately evaluated when they are more distinct from one another (more incommensurate in their terminology). For example, a gift-with-purchase promotion combines the first element (cash price of the item) with an incommensurate second element (merchandise gift item), whereas a price discount combines the first element (cash price of the item) with a commensurate second element (discount coupon for another purchase). The separate evaluations induced by the incommensurate elements create a more favorable evaluation over a monetarily equivalent promotion comprised of commensurate elements.

Applied to our context, our first element is the base compensation, which is always cash. Our second element is incentive compensation, which may be cash or merchandise. A cash incentive is commensurate with the first element, whereas a merchandise incentive is incommensurate. The latter encourages separate evaluations and greater perceived utility. In sum, a baseline cash plus

merchandise incentive plan should elicit more effort than the same baseline cash plus a monetarily equivalent cash incentive.

Preview of research and findings

Our primary goal is to uncover salesperson preferences across cash versus merchandise elements of an incentive plan. In doing so we hope to provide a rationale for the wide-spread popularity of such programs. To this end, we implemented a large-scale field intervention involving 580 salespeople of a division of a Midwestern frozen foods manufacturer. Their compensation plan consisted of a base commission for sales, plus bonuses for attaining (and exceeding) monthly targets. These bonuses were a combination of cash and "points"; the latter could be redeemed for a variety of merchandise from a catalog managed by a third party vendor. The existing plan would thus be described as one with merchandise incentives.

We switched the salespeople to a plan with all-cash bonuses, where the new bonuses are equal in their monetary cost to the firm as the old bonuses. With the firm's assistance, we assembled salesperson-level sales revenues over a 9-month spell, including both pre- and post- intervention months. From a series of statistical models, controlling for salesperson fixed effects, seasonality effects, year effects, and target effects, we concluded that our the switch from merchandise incentives to all cash reduced sales by 4.36%, on average. Exploiting the panel structure of our data to investigate causal effect heterogeneity, we estimated individual-level intervention effects. These estimates reveal a negative intervention effect among all the salespeople, but the size of the effect varies considerably, with the size of the sales drop increasing as we move from the lowest performing salespeople to the top performers.

To validate our posited causal chain (incentive plan $\Rightarrow$ effort $\Rightarrow$ sales), we estimated individual-level effort with a procedure first suggested by Bandiera et al. (2005). We find effort dropped from the pre- to post-intervention periods, and the individual-level effort drops parallel the individual-level sales drops, which affirms effort as the intervening variable.

Next, to uncover personal and social factors behind the heterogeneity in outcomes, we con-

ducted a post-study survey to measure several variables that might be plausibly tied to the use of separate mental accounts for cash and merchandise incentives. Using the survey responses to predict the previously estimated individual-level causal effects and effort changes, we find that both individual difference variables and social factors are influential. Specifically, salespeople from households with more discretionary income, and who employ a less abstract approach to future consumption plans (i.e., they thought more about money as the reward in itself rather than as purchases of goods and services enabled by the money) exhibited a greater sales drop as well as effort drop. Put differently, these salespeople exhibit a greater preference for merchandise incentives over cash incentives.

Finally, in an effort to ground our results in the prior literature we reviewed on stated preferences, we undertook a conjoint study to assess the stated preference for merchandise versus cash incentives among these salespeople. Consistent with prior work, our salespeople state that they get greater utility from cash incentives over monetarily equivalent merchandise incentives.

Overall, our work adds to the very limited evidence about behavioral reactions to merchandise versus cash incentives. It is noteworthy that merchandise incentives induced greater effort response to the tune of several millions of dollars compared to an equally costly cash-only incentive plan at this firm. This result goes a long way to explain the pervasive use of merchandise incentives in practice.

We now turn to describing the design of the field intervention, along with details of the institutional context. This is followed by a discussion of the results, and an estimation of individual-level efforts. This is followed by the post-intervention survey analysis and a final conjoint study to examine preferences across cash and merchandise incentives.

Design and Data

We collaborated on this study with a Midwestern consumer packaged goods manufacturer, whose revenue comes largely from frozen foods. It sells its brands nationally in grocery stores, including

big-box retailers, traditional chains, and independents. It utilizes a Direct Store Delivery (DSD) channel, and its route[4] salespeople play a key role within this setup.

Each salesperson is assigned to a number of retail stores (10, on average) that constitute a route, which is designed to maintain a minimum sales potential and to minimize travel distance to warehouses. Each route contains stores of varying sizes, and these routes are altered for reasons largely involving store openings or closings. Table 1 displays sample descriptive statistics. The average monthly potential for a route is roughly $\$101,114$. We have 12 months of pre-study data on monthly revenues and category sales for each route. Although these data are not at the salesperson level and contain no detailed compensation plan information, they supplement our before-after design, and help us better control history, seasonality, ability, route and other time invariant sources of heterogeneity.

Salespeople are required to visit each store at least once a week; on average, they visit each store three times a week. Their main tasks are to take orders, manage stock and execute sales promotion programs. They do not set wholesale prices or promotions budgets, but are responsible for store-level implementation of programs. These salespersons and the store managers have considerable leverage on what finally gets implemented in stores.

Figure 1 describes the original and revised compensation plans; in what follows we denote these plans as the "mixed" and "pure cash" regimes respectively. A few things are worthy of note. Salespeople are paid solely on output – there is no fixed wage. They receive commissions on sales and lump-sum bonuses for meeting their individual monthly targets. The firm generates these targets using an algorithm that incorporates prior year company sales and current year company sales goals, with adjustments for seasonality, store closing, openings, etc. Each route's monthly targets are announced at the beginning of the fiscal year and are not changed in the course of the year. Importantly, there was no change to previously announced targets following our intervention. The original monthly incentive payouts were a combination of cash and merchandise incentives, and amounted to about 11% of total monthly compensation cost for a salesperson meeting the target.

Our core data include sales revenues, compensation payouts, and sales targets for each route from July 2011 to March 2012, with the switch occurring in October 2011. We started with 676 routes, but 96 of these experienced some turnover during our observation period. In order to rule out validity problems,we excluded these 96 routes from our analysis, resulting in 580 unique observations in our estimation sample. Note that each route has the same salesperson throughout, which makes it easy to control for ability and other differences with salesperson fixed effects. [5]

By way of additional institutional background, the intervention occurred following a change in management; the sales executive in charge was promoted to a new position. Given the high costs of running the merchandise incentive program for the firm, the new manager wanted to understand its value more rigorously. We had an existing working relationship with the new manager, based on which he asked us to design a program to address his issue. It is important to note that the change was not spurred by some changes in competitive conditions, tax laws or demand.

We took the following things into consideration in implementing our pure cash regime. First, we controlled the timing of the change as well as the information provided to the salespeople. To mitigate demand artifacts, no advance notice or any specific reasons were given for the change, mirroring past practice. Second, our experimental design is a before-after design with repeated pre- and post-treatment observation periods, creating an interrupted time series panel. Unlike a treatment-control group design, there is no control group, which deserves some comment.

Treatment-control designs are not practical with a within-firm experiment (see Bandiera et al., 2011 for an extended discussion) on account of information about treatment leaking across groups, inviting inter-group comparisons, and injecting reactivity, demoralization and "John-Henry" style threats to internal validity (Shadish et al., 2002). Treatment-control designs are more practical within multi-firm settings, but multiple firm settings have their own disadvantages beyond the matter of getting cooperation, including heterogeneous (non-equivalent) treatment implementations, and irrelevant variation in products, prices, etc. That said, our conclusions must be tempered by the absence of a control group design - it is fair to say that we cannot make as strong a causal inference statement as might be possible with a treatment-control design.

Table 2 describes details of the mixed and pure cash regimes. Points were redeemable for merchandise on a site that resembled an online retail site such as Amazon.com. Items ranged from small-ticket items like music downloads to big-ticket items like televisions and vacation packages. The distribution and reimbursement of these points and merchandise was contracted out to a vendor; individual redemptions were not disclosed to the firm (or to us) to protect salesperson confidentiality. Table 3 displays the top five categories of redeemed merchandise.

The firm uses $3.75 per point in its internal accounting of the points program. Some institutional details are worth noting. The firm gave us access to the points program site, and we checked the merchandise price implied in dollars (i.e., points converted at the $3.75 rate) against the dollar price at Amazon.com for similar items. Invariably, we found lower prices at Amazon.com. For example, the number of points to redeem (at the 3.75 rate) a 40" television set was more than the equivalent Amazon price. There was no communication from the firm to the salesperson about this conversion rate, but a diligent individual could infer it from his own tax documents if so inclined.

Intervention Effects

Unadjusted Outcomes. Territory monthly sales decreased from $101,687 to $100,828.[6]. The pre- versus post-intervention sales distributions are shown in Figure 2, which plots kernel densities. There is a discernible shift to the left from pre- to post-intervention. Of course, given seasonality effects, and other factors that affect sales, this graphical shift is not sufficient to evaluate the intervention effect; we therefore turn to the statistical analyses below.

Intervention Effect. We estimate the treatment effect on monthly sales in a territory controlling for three sets of variables. First, differences across salespeople are controlled with a set of salesperson dummy variables. Second, annual patterns are contolled with year dummy variables, while monthly patterns are controlled with prior year territory sales for the same month. Third, differences across territories with respect to their changing market potential are controlled with the territory targets set for that month. Equation 1 is our estimation model, which is estimated by OLS with clustered territory-level errors.

$$(1) \qquad Sales_{it} = \alpha_i + \beta_2 Regime_t + \beta_3 Year_t + \beta_4 MonthSalesPriorYear_{it} + \beta_5 Target_{it} + \varepsilon_{it}$$

where $Sales_{it}$ are the sales of salesperson i in month t, α_i are the salesperson fixed effects, $Regime_t$ is the treatment dummy (set to 0 for mixed, 1 for pure cash), $Year_t$ are the year dummies, $MonthSalesPriorYear_{it}$ is monthly sales in prior year, $Target_{it}$ is the target set for salesperson i for month t, and ε_{it} are the clustered errors.

Table 4 reports the correlation matrix. Table 5 reports our estimates for three different specifications that are increasingly more comprehensive versions of Equation 1. The simplest specification in Column 1 includes only the *Regime*, *Year, MonthSalesPriorYear*, and *Target* variables. The intervention effect is significantly negative, ($\beta_2 = -4,873.40, p < .01$). Column 2 adds the α_i salesperson effects, which improves the model greatly. The intervention effect remains similar ($\beta_2 = -5,018.88, p < .01$); this corresponds to a mean sales drop of 4.36% arising from our intervention. This model is the specification we rely on for our subsequent discussion. The additional specifications in Table 5 provide robustness checks on our intervention effect size. Column 3 adds quadratic and cubic *Target* variables, yielding a productivity drop estimate of 4.53%.

Validity Threats and Robustness Checks

We consider a number of of nuisance explanations and possible threats to the validity of our statistical conclusions above. For brevity, we relegate the details to Web Appendix A. In particular, with our before-after design, it is important to rule out validity threats that might arise in the absence of a control group design.

Price Effects. Might sales have declined simply because an increase in price levels of the products? While we do not have the transaction-level data to build up a price index variable. Instead, we construct a firm-level price index by dividing the (nine) observed monthly firm-level revenue numbers into corresponding observed firm-level quantities. Notice, however, our depen-

dent variable is sales revenues, so our constructed price index introduces the same variable into the right hand side, making inferences problematic. With this caution in mind, for robustness, we estimated our intervention effect model (Equation 1) including this firm-level price index variable. The direction and significance of the intervention effect is unchanged, but the magnitude of the sales drop is now 6.90% (see Table WA.4).[7]

Tax Incentives. Might sales have declined because bonus points a) confer tax benefits to salespersons as compared to cash (e.g., might bonus points may be tax-advantaged like health or child-care benefits), and/or b) allow salespersons to strategically time their redemption to lower their tax burden? Two facts suggest that this is not the case.

First, the firm is required to pay the same amount of taxes on both cash incentives and points. Bonus points are converted to cash equivalent income at a fair market value – this converted amount is used for making withholding decisions and is included in the overall income on tax documents (e.g., W2 forms) for that salesperson. Second, the firm does not delay tax actions for points-based income until such time as the salesperson chooses to redeem points for merchandise. Withholding and reporting occur as soon as bonuses are earned. Neither cash nor points bonuses are ever postponed beyond the month in question. This ensures that while salespersons definitely have the freedom to decide when to redeem their points, this decision has no bearing on when the tax on the points earned is due. In summary, the tax treatment of all the components of compensation (salary, commission, points, etc.) is exactly similar and payable in that year's tax filing, ruling out tax incentives as a possible explanation for our results.

Category Trends. Might sales have declined because category sales dropped coincident with our intervention? Figure 3 plots year-over-year changes in firm and category sales over an extended period (January 2011 to March 2012) around the October 2011 intervention month. Observe that firm and category sales do not move in sync; in fact, they move in opposite directions in the months immediately following the intervention, so coincident category trends are ruled out as a possible confound.

Ratcheting. Might sales have decreased because the firm ratcheted targets upwards coinci-

dent with the intervention? Prior work has shown that ratcheting effects, if present, could lead to strategic gaming of effort. Might sales have decreased because the firm ratcheted targets upwards coincident with the intervention? Following Misra and Nair (2011; Table 2, pp. 244) and Chung et al. (2014; Table 3 pp. 169), we assess ratcheting by regressing territory targets against prior period sales and prior targets. If prior period sales influence the current target, this points to ratcheting. Table WA.1 shows the effect of s_{t-1} (sales in the prior period) on the target in the current month is insignificant ($0.15, p > 0.1$), but that the prior period quota ($Target_{(t-1)}$) has a positive, significant effect ($1.12, p < 0.01$). Thus, target updating is not driven by past sales, which rules out ratcheting as an alternative explanation. Further proof of robustness is provided by the fact that we control non-linearly for targets in our main regression.

Reactivity. Might sales have decreased following the intervention because salespeople reacted to the change *per se*, instead of the particular change in incentives from points to cash? We reasoned that reactance should be greater among salespeople with larger accumulated points balances. Table WA.2 shows that neither accumulated points nor its interaction with the regime dummy variable have a significant effect on sales, thus ruling out reactance arising from accumulated points.

Points Thresholds. Might sales have decreased because salespeople with cumulative point balances just below a certain level required to redeem a long anticipated item reacted negatively to the switch? By way of analogy, consider an airline customer with 24,000 miles who now places great value on the next 1,000 miles because of the 25,000 mile threshold for a set of rewards (see Drèze and Nunes, 2004 for evidence). We rule out this threat in two ways. First, redeemable merchandise items are available at various points level, starting from low thresholds. Second, we assembled cash values and points required for a selection of items from the program catalog, ranging from CDs to TVs. Figure 4 shows the calculated marginal value of a point (defined as the ratio of points in cash equivalence required to redeem the reward and the actual cash value of a reward) at different points totals. There is no visible increase in the marginal value of points, unlike the silver-gold-platinum thresholds common to airline programs.

Concurrent Historical Events. Might concurrent changes in the firm's marketing programs

have caused the sales drop? We checked with management, and verified that no new marketing program was initiated, nor was any pre-planned program discontinued during this period. Indeed, we chose the intervention period based in part on this issue.

Selection/sorting. Might our switch have prompted salespeople to leave, or others to join, changing the composition of the work force? Such effects can be quite large. For instance, Lazear's (2000) analysis of a change from fixed wages to wages plus commissions found that quits and entry coincident with the intervention had an effect on productivity equal to the effect of the incentive change itself. Is it possible that our productivity drop arose from high performers quitting or entering? We think this is unlikely given the features of our research design. First, we designed our intervention to minimize sorting. Knowing that a change in base compensation would induce quits and entry (see Lo, Ghosh and Lafontaine, 2011), we elected to maintain the prior base wage.

Additionally, recall that 96 territories experienced some type of turnover during our observation period (quits, entry, or route changes) so we excluded them from further analyses. As a robustness check, we re-estimated our intervention model with these territories included. Table WA.3 reports estimates of the intervention effect that are identical in direction and significance, and close in magnitude as well, which rules out selection/sorting as a possible confound.

Sales Equation Specification. Might our results be an artefact of the linear specification of our sales equation (Equation 1)? Prior work has alternatively used a Cobb-Douglas function (e.g., Bandiera et al. 2005). To verify the robustness of our linear specification, we estimated a Cobb-Douglas specification of Equation 1. The coefficients reported in Table WA.5 show no material changes in results. [8]

Summing up, it is reasonable to say that the finding of a significant drop in sales can be causally attributed to the shift from mixed to pure cash incentives. Before we proceed to examine heterogeneity around this average effect, we take a detour to examine the issue of stated versus revealed preferences of salespeople.

Self-Reported Preferences. Thus far, all our analyses have focused on the salespeople's behavioral responses, i.e., sales changes. In standard terminology, we are operating in the realm of the

salesperson's revealed preferences. What about their stated preferences? Do they express a stated preference for merchandise over cash consistent with their behavioral responses? Or do they express a stated preference for cash as consistently reported in the literature (e.g., Churchill, Ford and Walker, 1985; Kube et al., 2012)? To ground our work in this literature, we conducted an on-line conjoint study of these salespeople following the experimental period.

We utilized a choice-based conjoint design, where a participant is asked to choose from hypothetical sets of incentive plans (see Web Appendix C). Each plan within a choice set is a different combination of cash and merchandise points. We chose these levels of cash and points to be within the realistic range of their previous plan. In all, 148 salespeople responded to our conjoint study. Table 6 reports the coefficients of their utility functions estimated as a discrete choice conditional logit model fitted to their responses.

From the utility coefficients for cash (4.06, $p < .01$; 7.96, $p < .01$; 9.33, $p < .01$), we see that more money is preferred to less money. Likewise, more points are preferred to fewer points (2.04, $p < .01$; 3.87, $p < .01$; 6.57, $p < .01$). More importantly, the coefficients allow us to compare the ratio of the utility of an additional dollar paid out as cash versus that same dollar's utility when it is converted to merchandise points (at the firm's cost of \$3.75 per point) and paid out. Performing these calculations on the coefficients in Table 6 reveals that our salespeople self-report more utility from a cash payout of a dollar versus that same dollar paid out in points. For example, each additional plan dollar between the \$10 and \$20 levels yields 0.39 utiles ($(7.96-4.06)/10 = .39$), while each additional point between the 20 and 30 points levels yields only 0.18 utiles ($(3.87-2.04)/10 = 0.18$), even though the latter costs the firm \$3.75. Our result replicates prior stated preference studies, but it also documents the striking divergence of self-reports from behavioral reports. This divergence has hitherto escaped verification on account of the lack of any multi-method study undertaken previously. Plainly, self-reports of incentive preferences have limited value to scholars and managers.

Heterogeneous Intervention Effects

Thus far, our analyses show that average sales decreased significantly following our intervention; this result is robust to a number of alternative explanations and survives a set of robustness checks. However, our constant treatment effect analysis obscures potentially heterogeneous effect sizes, which is hinted at in Figure 2. Notice that the visual gaps between the pre- and post-intervention sales distributions are unequal across the range of their values. Heterogeneous causal effects have important managerial consequences. For example, our significant negative effect might be comprised of a small number of large negative individual-level effects outweighing larger numbers of smaller, positive effects, which implies very different personnel policies from more uniformly negative, albeit smaller individual-level effects. We examine the heterogeneity of treatment effects through a random coefficients specification below.

Random Coefficients. We estimate the following random coefficients model, identical to our previous fixed coefficients model (equation 1), except that we now have randomly varying coefficients for salespeople and regime.

$$(2) \qquad Sales_{it} = \alpha_i + \beta_{2i} Regime_t + \beta_3 Year_t + \beta_4 MonthSalesPriorYear_{it} + \beta_5 Target_{it} + \varepsilon_{it}$$

where $\alpha_i \sim \mathrm{N}(\alpha, \sigma_\alpha)$ are normally distributed salesperson effects, $\beta_{2i} \sim \mathrm{N}(\beta_2, \sigma_\beta)$ are normally distributed regime effects, and the other coefficients are as described previously for the fixed effects regression. Table 7 reports the estimates of Equation 2.

The salespeople and regime effects estimates are both significant, and there is significant heterogeneity in each case. Figure 5 plots individual-level posterior estimates of the regime effect. There is a broad range $[-8,090.67; -2,185.78]$ of regime effects around the median (-4,909.26) with all salespeople exhibiting a negative regime effect.

Effort as Intervening Variable

Our analysis so far shows that sales decreased with our intervention, and this effect varied across salespersons. From a theoretical standpoint, compensation induces effort, which yields sales. To the extent that other factors are controlled for, it is inevitable that one will find a reduction in effort accompanying a sales decline. Nevertheless, individual estimation of effort is essential to both validate this causal chain, as well as uncover possible heterogeneity across salespeople. Unlike a laboratory study, however, direct measures of effort are neither feasible nor desirable in our field experiment, whose strength is that our salespeople are unaware of the experiment (Harrison and List, 2004). Querying them about effort would introduce demand artifacts. List and Rasul (2011) observe that very few field experiments provide evidence about the causal chain, and recommend structural estimation to recover these intervening variables. As such, we follow Bandiera et al. (2005) to recover individual salesperson effort for each time period. Intuitively, after controlling for seasonality, etc., observed sales are a function of salesperson ability and effort. Thus, if we were to observe a salesperson over time, we can isolate the ability contribution through a fixed effect for each regime. Recall that ability remains constant, so changes in an individual's estimated effects across the two regimes derives entirely from changes in effort. We implement their procedure as follows.

First, we re-estimate the sales regression (equation 1), but with an added variable, viz., the interaction of the salesperson fixed effect with the regime dummy. Note that all continuous variables are in logarithms in the specification below

(3)

$$Sales_{it} = \alpha_i + \beta_2 Regime_t + \beta_3 Year_t + \beta_4 MonthSalesPriorYear_{it} + \beta_5 Target_{it} + \beta_6 [Regime_t \cdot \alpha_i] + \varepsilon_{it}$$

Next, we compute salesperson i's effort during period t within regime $\{s \in mixed, pure\}$ by adding the estimate of that salesperson's effect to his residual from the regression equation above.

(4)

$$\hat{e}^s_{it} = \hat{\alpha}^s_i + \varepsilon^s_{it}$$

where $\hat{\alpha}_i^s$, the ith salesperson effect during regime s, is given by $\hat{\alpha}_i^s = \alpha_i + \beta_2 Regime_t$, and ε_{it}^s is that part of the salesperson's productivity not captured by the observables included in equation 3. This procedure yields an estimate of each salesperson's effort for every month t under incentive scheme s.

Figure 6 plots the distribution of these effort estimates. The effort distribution is visually shifted to the left in the pure cash regime. The mean effort level across salespersons and time periods drops from 0.58 in the mixed regime to 0.54 in the pure cash regime. Of the 579 salespersons, a large majority (370; 63.90%) reduced their effort (10.70%, on average), while 209 salespeople (36.10%) increased their effort (7.45%, on average). These differences are highly significant ($t = -8.37; p < .01$).[9]

Consider the constituent components of $\hat{e}_{it}^s$ across regimes; recall the first part comes from a change in the fixed effect across regimes, while the second part comes from the residual change across regimes. We would hope that our estimates are largely driven by the first component. To this end, we plotted the residuals of Equation 3 for the observations from the two regimes in Figure 7. The mean residual is the same for the pre- and post-intervention observations, while its variance drops slightly, suggesting that our effort estimates from Equation 4 are almost totally attributable to the salesperson effect.

Productivity Shifts. Are there discernible patterns in our effort estimates across salespeople; in particular, is relative sales productivity related to effort change? Do more productive salespeople show a bigger drop in effort across regimes than their productive counterparts? This is a variation of the sorting effect, except that instead of quits and hires, it asks whether the rank ordering of treated salespeople changes across incentive regimes. To this end, we compute the rank correlation between observed salesperson productivity and their estimated effort change. The negative correlation (-0.10,p = 0.00) suggests that more productive salespeople had significantly greater reductions in in effort.

To sum up the analyses above, our intervention lowered productivity through lower induced effort. Also, the effort reductions were larger amongst more productive salespeople, leading to

disproportionately greater sales reductions from those salespeople that the firm values the most, i.e., the highest productivity types.

Sources of Heterogeneity

We have established that the regime effects and effort changes are heterogeneous across salespeople, which suggests that their reliance on mental accounts also varied. To better understand the heterogeneity in the use of mental accounts, we solicited salespersons' participation in an online survey. Identifying ourselves as academic researchers, we assured them their individual responses would not be identified or shared with the firm.

Starting with the global wealth-maximizing agent as a straw man, we examined the broad literature on contextual preferences to find demographic and psychological factors that might explain separate accounts for cash and merchandise. Armed with a list of such factors, we developed a short questionnaire that measured a number of speculative variables. Following a pretest with the firm's sales managers to verify wording, response formats and clarity of instructions, we made changes based on their feedback, and distributed the surveys electronically.[10] After two reminder emails, we received 263 completed questionnaires. Of these, 47 were eliminated due to excessive missing data, leaving a final sample of 216. To assess non-response bias, we compared respondents to non-respondents along known variables (sales, etc.) and found no statistically significant difference between the two groups.

Household Resources. Our survey asked for the average fraction of monthly income spent on discretionary expenditures (entertainment, restaurants, gifts, movies, electronics, etc.). Some work (Ashraf et al. 2006; Laibson, 2001) suggests that people are willing to purchase commitment devices to mitigate self-acknowledged difficulties in managing consumption plans over time. Since points can only be redeemed for merchandise from the catalog, it might be easier to set these points aside, and to buffer them from other demands on discretionary household resources. In effect, points might serve as a superior commitment device compared to cash. The availability of slack resources should impact the value of merchandise incentives as a commitment device - the

more the slack resources an individual possesses, the greater the need for commitment devices. To capture these effects we constructed the *fraction spent on non-discretionary expenses*. As this non-discretionary fraction increases, slack diminishes and so does the need for commitment. Thus, merchandise incentives should be less attractive.

Redemption Frequency. Our survey asked how often they redeemed points. This builds on the same idea of commitment. Commitment devices should be more valuable for planning larger, more infrequent purchases. *Redemption Frequency* should be lower for people who value commitment. Those who wait to accumulate more points value commitment more, so a lower frequency of redemption should be associated with a higher utility for merchandise incentives.

Household size. Our survey asked for the number of people in the household. A large literature suggests that multi-person purchases are determined through a variety of mechanisms such as bargaining, power, fairness, etc. that bring together diverse individual preferences (see Chiappori, 1988 for an economics-oriented review, or Arora and Allenby, 1999 for a marketing viewpoint). Cash rewards are more likely to be considered part of the joint income of the household and bargained away without providing additional utility to the individual. On the other hand, non-monetary rewards are less fungible and transferable, giving the salesperson greater control over their allocation. In concrete terms, an extra $500 in cash is much more likely to go into the household expenditure kitty directly, than points that can potentially buy an iPad worth $500. In other words, non-monetary rewards are less likely to be bargained away in any negotiating process, thus enhancing the salesperson's individual utility. This suggests that larger family size should be associated with a higher utility for merchandise incentives.

Abstraction Level. Individuals differ in their chronic *abstraction levels* used to represent and evaluate everyday objects and events (Dhar and Kim, 2007; Fiedler, 2007; Kardes et al., 2006; Trope and Liberman, 2000). As applied to our context, a salesperson might think about a cash incentive in a broad, abstract fashion (e.g., what could I buy with the money?) or in a more narrow, concrete fashion (e.g., exactly how much more money will I make?).

Our survey asked whether they thought about the uses of the funds or were more focused on

the dollars themselves. Higher numbers on a 7-point scale reflect more abstract thinking on our *Money Abstraction Level* measure. Salespeople who view cash more abstractly see the flexibility of cash, and hence might prefer cash to merchandise incentives.

Hedonic Consumption. Our survey asked salespeople for their last two redeemed items. These can be organized along a utilitarian-hedonic continuum, with utilitarian objects providing functional, instrumental, or practical benefits, and hedonic objects providing aesthetic, and experiential enjoyment (e.g., Holbrook, 1999; Batra and Ahtola, 1990; Dhar and Wertenbroch, 2000; Strahilevitz and Myers, 1998). Applied to our context, a utilitarian merchandise item is more readily compared on features, price, etc. to comparable items from the general marketplace, and hence is likely to be valued at its market price. Hedonic merchandise items are processed in a more affective fashion, making comparisons more difficult, which enhances their evaluation over the cash equivalent (e.g., Loewenstein et al. 2001). We suspect that salespeople with a preference for merchandise incentives are attracted to the more hedonic items. Hence, they should be more likely to redeem such items.

We trained two coders to independently categorize the reported items as either utilitarian or hedonic (Web Appendix C reports the instructions). Of 230 items, 73 items were categorized as hedonic and 157 as utilitarian. We code *Hedonic Redemption Level* as 1, if both of the redeemed items were utilitarian, 1.5 if one item were utilitarian and the other hedonic, and 2 if both items were hedonic. If the salesperson redeemed only one item, the *Hedonic Redemption Level* is 1 if the product is utilitarian and 2 if hedonic.

Results. We regressed the five measures described above on our estimates of a) individual-level regime effects and b) individual-level effort changes. Tables 8 and 9 report the results for the regime effect and effort change variables respectively. These dependent variables are coded such that higher values denote a smaller difference in preference for merchandise incentives over cash incentives. Keeping this in mind, we can interpret the significant coefficients; viz., household income fraction spent on non-discretionary expenses and abstraction level.

As the household's fraction of income spent on non-discretionary expenses increases, the sales-

person's preference for merchandise over cash shrinks ($\beta = 574.74,\ p < .05$ and $\beta = .05,\ p < .1$ in Tables 8 and 9 respectively). Evidently, as the household's slack resources shrink, commitment needs drop, and the salesperson's preference for merchandise over cash incentives shrink correspondingly. Second, as the salesperson's chronic predisposition towards money becomes more abstract, the preference for merchandise over cash shrinks ($\beta = 94.49,\ p < .05$ and $\beta = .01,\ \mathrm{p} < .01$ in Tables 8 and 9 respectively). Salespeople who think more about the multiple, functional uses of money have smaller preferences for merchandise over cash incentives. None of the other measures are significant. However, we caution these results are only suggestive because they are not stable when all the measured background variables are entered simultaneously.

Taken together, we see that an individual's social context (household slack) as well as an individual difference variable (abstraction level) appear to influence his use of separate accounts for merchandise and cash. Note that these results are robust against the usual endogeneity problem with cross-sectional survey data because our dependent variables come from our experimental intervention.

Conclusions

Our primary goal is to obtain insight into cash versus merchandise incentive plans. From a theory standpoint, the workhorse principal-agent model favors the flexibility of cash incentives, as do stated preference studies. Nevertheless, merchandise incentive plans are used widely in practice.

To understand the attractiveness of such plans, we designed and implemented a large scale field intervention at a national frozen food manufacturer. We changed the extant incentive package for 580 route salespeople from a cash plus merchandise incentive plan to an equivalent all-cash incentive plan. Equivalence was achieved by converting the points redeemable for merchandise into cash at the firm's dollar cost of creating a point and paying out the combined value in cash. The base wage was left unchanged.

Sales dropped by an estimated 4.36% after controlling for a variety of background factors.

Capitalizing on our panel design, we are able to estimate individual-level effort, which is our posited intervening variable between incentives and output. We find that pre-post intervention reductions in effort parallel the reductions in sales, which validates our causal chain. Estimating individual-level effects to address heterogeneity, we find that all our salespeople behaved as if they preferred merchandise incentives over equivalent cash; strikingly, the most productive salespeople displayed the largest preference for merchandise.

Our results are best explained by theories of separate mental accounts for merchandise and cash. In this our explanation is similar to recent work in a very different domain, namely the stockpiling of loyalty points. In that work, Stourm et al. (2015) pin down mental accounting as the theoretical rationale for this seemingly non-rational behavior. To gain further insight into this explanation, we identified possible correlates of separate accounts. Using measures from our post-experiment survey, we find some suggestive results. Salespeople from households with greater discretionary income prefer merchandise over equivalent cash to a greater degree. Similarly, salespeople who perceive cash in a more concrete fashion also preferred merchandise incentives over cash incentives to a greater degree. We caution, however, that these are not robust to multi-collinearity.

Summing up, our study establishes quite robustly that cash is not king in compensation plans. Our results provide a business case for merchandise incentive plans; they work well because salespeople behave as if they prefer merchandise incentives over equally costly cash incentives. However, despite the significant covariates of their preferences from our survey, it is premature to offer a comprehensive model of these drivers of separate preferences for cash and non-cash compensation payout.

We contribute to sales compensation method, theory and practice. First, our multi-method design shows up the fragility of stated preferences for cash incentives over merchandise incentives, making conjoint studies a hazardous basis for choosing compensation plans. Substantively, our large-scale field study adds to the very limited evidence about behavioral reactions to merchandise versus cash incentives. The former type of incentive induced greater effort response to the tune of several millions of dollars compared to an equally costly cash-only incentive plan at this firm. This

result explains the pervasive use of merchandise incentives in practice. From a theory standpoint, the individual-level heterogeneity analyses afforded by our repeated measures design pins down effort as the intervening variable between incentives and outcomes. Also, our post-study survey points to abstraction level, and financial slack as background factors that explain the greater use of separate mental accounts.

We close with some comments on future research. First, we re-iterate that our before-after design leaves history as a viable alternative explanation, despite all our robustness analyses to rule out a variety of specific threats. Clearly, employing designs with clear control groups would be a big step forward. Conceptually, a promising, unaddressed issue is the matter of including symbolic rewards into the mix. Are the three types of rewards (cash, merchandise, and symbolic incentives) additive, complementary, or do they "crowd out" each other? Finally, we have not solved for the optimal plan involving multiple types of payouts. This is an important and computationally laborious task that we leave for future research.

References

Akerlof, G. (1982), "Labour contracts as a partial gift exchange." *Quarterly Journal of Economics*, 97, 543–569.

Arora, Neeraj and Allenby, Greg M (1999), "Measuring the influence of individual preference structures in group decision making." *Journal of Marketing Research*, 36, 476–487.

Ashraf, N, Karlan, Dean, and Yin, Wesley (2006), "Tying odysseus to the mast: Evidence from a commitment savings product in the Philippines." *The Quarterly Journal of Economics*, 121, 635–672.

Bandiera, O, Rasul, I., and Bar (2011), "Field experiments with firms." *Journal of Economic Perspectives,* 25 (3) 63-82.

Bandiera, O, Rasul, Imran, and Barankay, Iwan (2005), "Social preferences and the response to incentives: Evidence from personnel data." *Quarterly Journal of Economics*, 120, 917–962.

Batra, R. and Ahtola, O. T. (1990), "Measuring the hedonic and utilitarian sources of consumer attitudes." *Marketing Letters*, 2 (2), 159 –70.

Chiappori, P. (1988), "Nash-bargained household decisions." *International Economic Review*, 32, 791–96.

Chung, Doug J., Thomas Steenburgh, and K.Sudhir (2014), "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans." *Marketing Science*, 33(2), 165 –187.

Churchill, G.A., N.M., Ford, and Walker, O.C. (1985), "Differences in the attractiveness of alternative rewards among industrial salespeople: Additional evidence." *Journal of Business Research*, 13, 123–128.

Dhar, R. and Wertenbroch, K. (2000), "Consumer choice between hedonic and utilitarian goods." *Journal of Marketing Research*, 37, 60 –71.

Dhar, Ravi and Kim, Eunice Y. (2007), "Seeing the forest or the trees: Implications of construal level theory for consumer choice." *Journal of Consumer Psychology*, 17 (2).

Dorotic, Matilda, Bijmolt, Tammo H.A., and Verhoef, Peter C. (2012), "Loyalty programmes: Current knowledge and research directions." *International Journal of Management Reviews*, 14, 217–237.

Drèze, X. and Nunes, J.C. (2004), "Using combined-currency prices to lower consumers' perceived cost." *Journal of Marketing Research*, 41(February), 59–72.

Fiedler, Klaus (2007), "Construal level theory as an integrative framework for behavioral decision- making research and consumer psychology." *Journal of Consumer Psychology*, 17 (2).

Gneezy, Uri and List, John A (2006), "Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments." *Econometrica*, 74 (5), 1365–1384.

Gravalos and Jack, John M. (2007), "The trouble with money." *Business Insider Report*.

Harrison, Glen W. and List, John (2004), "Field experiments." *Journal of Economic Literature*, 42, 1009–1055.

Holbrook, M. B. (1999), "Introduction to consumer value." *Consumer Value: A Framework of Analysis and Research*, 1 –28.

Incentive Federation (2013), "A study conducted among current user of merchandise and travel items for motivation/incentive applications." Westfield, NJ.

Kahneman, D. and Tversky, A. (1979), "Prospect theory: An analysis of decision under risk." *Econometrica*, 47,263 –91.

Kardes, Frank R., Cronley, Maria L., and Kim, John (2006), "Construal-level effects on preference stability, preference-behavior correspondence, and the suppression of competing brands." *Journal of Consumer Psychology*, 16(2).

Keh, H.T. and Lee, Y.H. (2006), "Do reward programs build loyalty for services? the moderating effect of satisfaction on type and timing of rewards." *Journal of Retailing*, 82, 127-136.

Kivetz, R. (2003), "The effects of effort and intrinsic motivation on risky choice." *Marketing Science*, 22, 477-502.

Kivetz, R. (2005), "Promotion reactance: the role of effort-reward congruity." *Journal of Consumer Research*, 31, 725-736.

Kube, Sebastian, Maréchal, Michel André, and Puppe, Clemens (2012), "The currency of reciprocity: Gift exchange in the workplace." *The American Economic Review*, 102 (4).

Laibson, David (2001), "A cue-theory of consumption." *The Quarterly Journal of Economics*, 116, 81–119.

Lazear, Edward P. (1986), "Salaries and piece rates." *Journal of Business*, 59, 405–431.

Lazear, Edward P. (2000), "The power of incentives." *The American Economic Review*, 90, 410–414.

List, John and Rasul, Imran (2011), *"*Field Experiments in Labor Economics*" in Handbook of labor economics*, 104–228. Elsevier.

Lo, Desmond (Ho-Fu), Ghosh, Mrinal, and Lafontaine, Francine (2011), "The incentive and selection roles of sales force compensation contracts." *Journal of Marketing Research*, 48, 781–798.

Loewenstein, G., Weber, E. U., Hsee, C. K., and Welch, N. (2001), "Risk as feelings." *Psychological Bulletin*, 127, 267 –86.

Marino, Anthony M. and Zabojnik, Jan (2008), "Work-related perks, agency problems, and optimal incentive contracts." *RAND Journal of Economics*, 39(2), 565 –585.

Melancon, J. Phillips, Noble, S.M., and Noble, C.H. (2011), "Managing rewards to enhance relational worth." *Journal of the Academy of Marketing Science*, 39, 1–22.

Misra, S. and Nair, H. (2011), "A dynamic structural model of sales-force compensation: Estimation and field implementation." *Quantitative Marketing and Economics*, 9(3), 211 –25.

Nunes, Joseph C. and Park, C. Whan (2003), "Incommensurate resources: Not just more of the same." *Journal of Marketing Research*, (40), 26–38.

Oyer, Paul (2008), "Salary or benefits?" *Research in Labor Economics*, 28, 429 –467.

Oyer, Paul and Schaefer, Scott (2005), "Why do some firms give stock options to all employees?: An empirical examination of alternative theories." *Journal of Financial Economics*, 76, 99 –133.

Rebitzer, J. and Taylor, L. (2011), "Extrinsic rewards and intrinsic motives: standard and behavioral approaches to agency and labor markets" in *Handbook of Labor Economics*, Amsterdam: North-Holland, 701– 772.

Shadish, William R., Cook, Thomas D., and Campbell, Donald T. (2002), *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin Company.

Shaffer, V.A and Arkes, H.R. (2009), "Preference reversals in the evaluation of cash versus noncash incentives." *Journal of Economic Psychology*, 30, 859 –872.

Smith, A. and Sparks, L. (2009), "Reward redemption behaviour in retail loyalty schemes." *British Journal of Management*, 20, 204-218.

Steenburgh, Thomas and Ahearne, Michael (2012), "Motivating salespeople: What really works." *Harvard Business Review*, 90, July-August, 70-75.

Stourm, Valeria, Eric T. Bradlow and Peter S. Fader (2015), "Stockpiling Points in Linear Loyalty Programs", Journal of Marketing Research, 52 (April), 253-267

Strahilevitz, M. and Myers, J. G. (1998), "Donations to charity as purchase incentive: How well they work may depend on what you are trying to sell." *Journal of Consumer Research*, 24, 434 –46.

Thaler, R. H. (1985), "Mental accounting and consumer choice." *Markeing Science,* 4, 199 –214.

Thaler, R. H. (1999), "Mental accounting matters." *Journal of Behavioral Decision Making*, 12, 183 –206.

Trope, Y. and Liberman, N. (2000), "Temporal construal and time-dependent changes in preference." *Journal of Personality and Social Psychology*, 79, 876 –89.

Yi, Y. and Jeon, H. (2003), "Effects of loyalty programs on value perception, program loyalty, and brand loyalty." *Journal of the Academy of Marketing Science*, 31, 229-240.

Notes

[1]These "merchandise" catalogs include goods like appliances, as well as services like vacations, movie downloads, sporting events, etc.; their common feature is that all these items carry a real economic cost to the firm.

[2]In certain instances, firms get discounts for purchasing items in bulk which might make these programs more appealing, but this is definitely not the case here.

[3]Principal-agent models, the workhorse of the academic literature on salesforce compensation, have been expanded recently to go beyond cash, and incorporate intangible, psychological considerations such as status, recognition, effort norms and reference wages (see Rebitzer and Taylor, 2011 for a review). Even these extensions speak to purely intangible payouts that impose zero economic cost to the firm; in contrast, in these models, merchandise incentive programs are folded into monetary compensation at some conversion rate, making them equivalent to cash in their impact on agents' effort choices.

[4]"Route", "territory" and "salesperson" are used interchangeably.

[5]Each salesperson is uniquely identified with each route, and with no changes in either person or route in our data, changes in route sales are uniquely attributed to changes in effort by each salesperson.

[6]Note that while we do have weekly data for some of the period, we perform our analysis at the monthly level because the frequency of the monthly data is much more in line with everything related to compensation (e.g., targets), so salesperson decision making is also sensibly studied at that level.

[7]In Web Appendix B, we replicate the analyses presented in this paper with the firm-level price variable included. The significance and direction of effects remain unchanged.

[8]Parenthetically, all the analyses have been replicated with the Cobb-Douglas form. The results do not change in direction or significance; these results are available on request.

[9]Web Appendix D details this analysis.

[10]The items are described in Web Appendix C.

Table 1: Observation Period Route Statistics (June 2011-March 2012)

	N of obs	Mean	Min	Max	Std. Dev
Monthly Sales ($)	5,220	101,114	3,246	277,837	101,115
Salesperson Tenure (years)	580	8.17	.33	31	
Monthly Targets ($)	5,220	114,596	33,693	401,255	114,596
Monthly Sales Prior Year ($)	5,220	113,933	28,261	352,588	113,933
New Hires			60		
Quits			75		
Routes = 676					

Table 2: Compensation Plan Structure

Time Period	Plan Elements	Description
July 2011-September 2011(Mixed Regime)	Commission	4% of credited monthly sales
	Lump-sum Bonus at Target	$300
	Bonuses exceeding target	$75 plus 20 points for exceeding target; $150 plus 40 points for exceeding 108% of target; $225 plus \60 points for exceeding 115% of target
October 2011-March 2012 (Pure-cash Regime)	Commission	4% of credited monthly sales
	Lump-sum Bonus	$300
	Bonuses exceeding target	$150 for exceeding target; $300 for exceeding 108% of target; $450 for exceeding 115% of target

Table 3: Redeemed Merchandise Items

Category	Sample Items
Tools and Appliances	Leaf blower, Household appliance, Small appliance, Grill, Coffee mug
DVDs, games, mp3s	Movie DVDs, PS3 & Xbox360 games, Songs (mp3)
Game controllers	PS3, Nintendo3DS, Wii Radio, Xbox 360 controllers
Computer, Accessories	Keyboard, Monitor, Speakers, Portable hard drive
Entertainment	TVs, iPod, Speaker, MP3 player, Radio

Table 4: Correlation Matrix

	Sales	Target	MonthSalesPriorYear
Monthly Sales	1.00		
Monthly Target	0.69	1.00	
MonthSalesPriorYear	0.60	0.71	1.00
N of Obs	5,220		

Table 5: Intervention Effect (Fixed Coefficients)

Dependent Variable: Territory Monthly Sales ($)			
Independent Variable	Model 1	Model 2	Model 3
Regime	-4,873.40(780.02)***	-5,018.88(688.15)***	-5,221.43(691.28)***
Target	0.51(0.01)***	0.16(0.01)***	0.65(0.02)***
$Target^2$	No	No	-3.28E-06(6.76E-07)***
$Target^3$	No	No	6.52E-12(1.21E-12)***
MonthSalesPriorYear	0.17(0.01)***	-0.06(0.01) ***	-0.06(0.01)***
Lagged Monthly Sales	No	No	No
Year Effects	Included	Included	Included
Salesperson Effects	No	Included	Included
Constant	22,020.03(1,247.49)***	47,376.69(6,849.5)***	22,802.95(8,242.56)***
Adj. R^2	0.50	0.66	0.66
N of Obs.	5,220	5,220	5,220
N of salespeople = 580			

Table 6: Conjoint Analysis (Discrete Choice)

Factor	Factor Level	Coefficient (Std. Err.)
Monetary Incentives		
	$10	4.06 (0.17) ***
	$20	7.96(0.38) ***
	$30	9.33 (0.48) ***
Points for Merchandise Incentives		
	10 points	2.04 (0.27) ***
	20 points	3.87 (0.30) ***
	30 points	6.57 (0.51) ***
R^2		0.48
N of Obs. = 2,682	# salespersons=148	

*** Significant at 99% level.We have 4 levels for monetary incentives (0, $10, $20, $30) and 4 levels for non-monetary incentives (0, 10, 20, 30).

Table 7: Intervention Effect (Random Coefficients)

Dependent Variable: Territory Monthly Sales ($)		
Fixed Part	Independent Variable	Coefficient(Std. Err.)
	Regime	-4,959.14(717.43) ***
	Target	0.36(0.01) ***
	MonthSalesPriorYear	0.08(0.01) ***
	Salesperson Effects	49,457.96(1,731.48) ***
	Year Effects	Included
Random Part		
	σ_β[Regime]	1,283.07(172.48)***
	σ_α[Salesperson Effect]	12,849.23(944.94)***
Adj. R^2		0.50
N of Obs. = 5,220	N of salespeople=580	*** significant at 99% level

Table 8: Sources of Individual-Level Regime Effect

Independent Variable	Coefficient (Std. Err.)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent Variable: Individual-Level Regime Effect						
Non-disc Exp. Fraction	574.74 **					622.26
	(287.94)					(379.33)
Redemption Frequency		20.84				52.17
		(66.31)				(94.79)
Household Size			-36.89			-77.75 *
			(35.54)			(42.50)
Abstraction Level				94.49 **		89.32
				(41.75)		(57.90)
Hedonic Redemption					-135.47	-124.04
					(252.63)	(252.11)
Constant	-5,566.93	-5,238.67	-5,069.42	-5,663.39	-5,097.81	-5,783.52 ***
	(193.90)***	(155.88)***	(139.67)***	(219.30)***	(369.96)***	(579.53)
R^2	0.01	0.00	0.01	0.02	0.00	0.06
N of Obs	216	216	216	216	216	216

*** Significant at 99% level;** Significant at 95% level; * Significant at 90% level.

Table 9: Sources of Individual-Level Effort Change

Independent Variable	Coefficient (Std. Err.)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent Variable: Individual-Level Effort Change						
Non-disc Income	0.05 *					0.09 *
	(0.03)					(0.04)
Redemption Freq		0.004				-0.001
		(0.007)				(0.008)
Household Size			0.005			0.002
			(0.004)			(0.004)
Abstraction Level				0.01 ***		0.01 **
				(0.004)		(0.005)
Hedonic Redemption					-0.01	-0.004
					(0.02)	(0.03)2
Constant	-0.10 ***	-0.09 ***	-0.10 ***	-0.13 ***	-0.06 *	-0.17 ***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.03)	(0.05)
R^2	0.0055	0.0022	0.0108	0.0254	0.0012	0.0632
N Obs	216	216	216	216	216	216

*** Significant at 99% level; ** Significant at 95% level; * Significant at 90% level.

Figure 1: COMPENSATION STRUCTURE

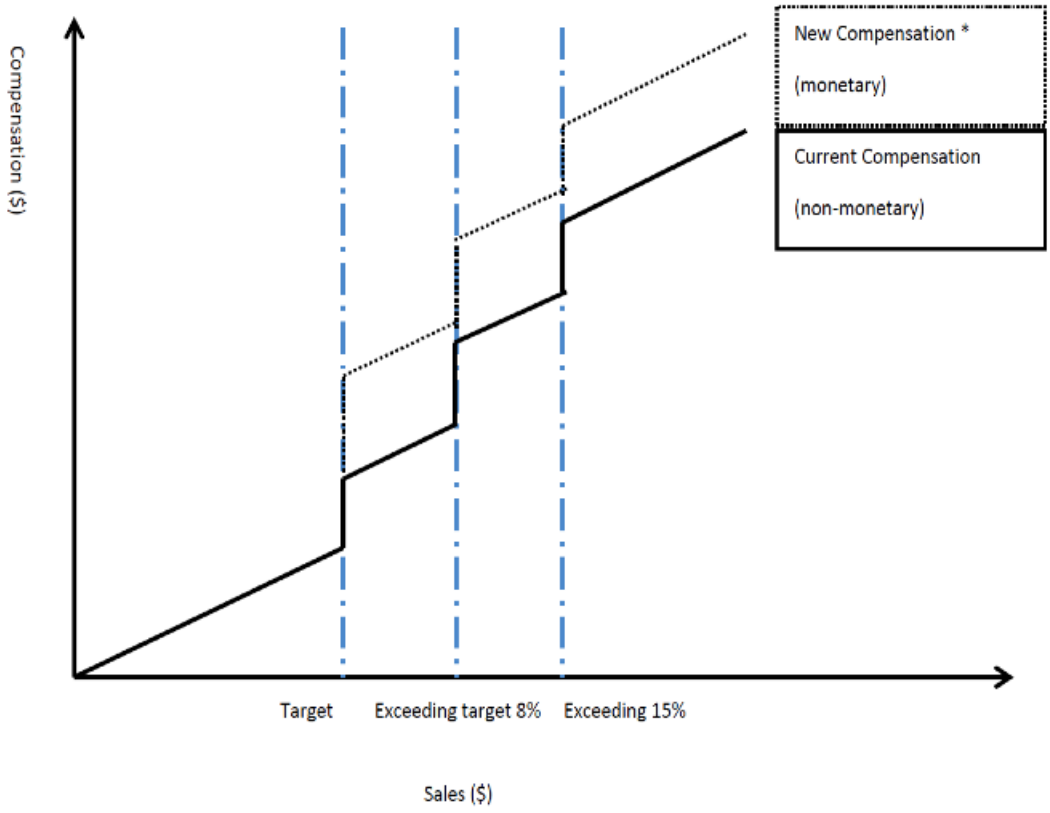


Figure 2: PRE-POST MONTHLY SALES DENSITY

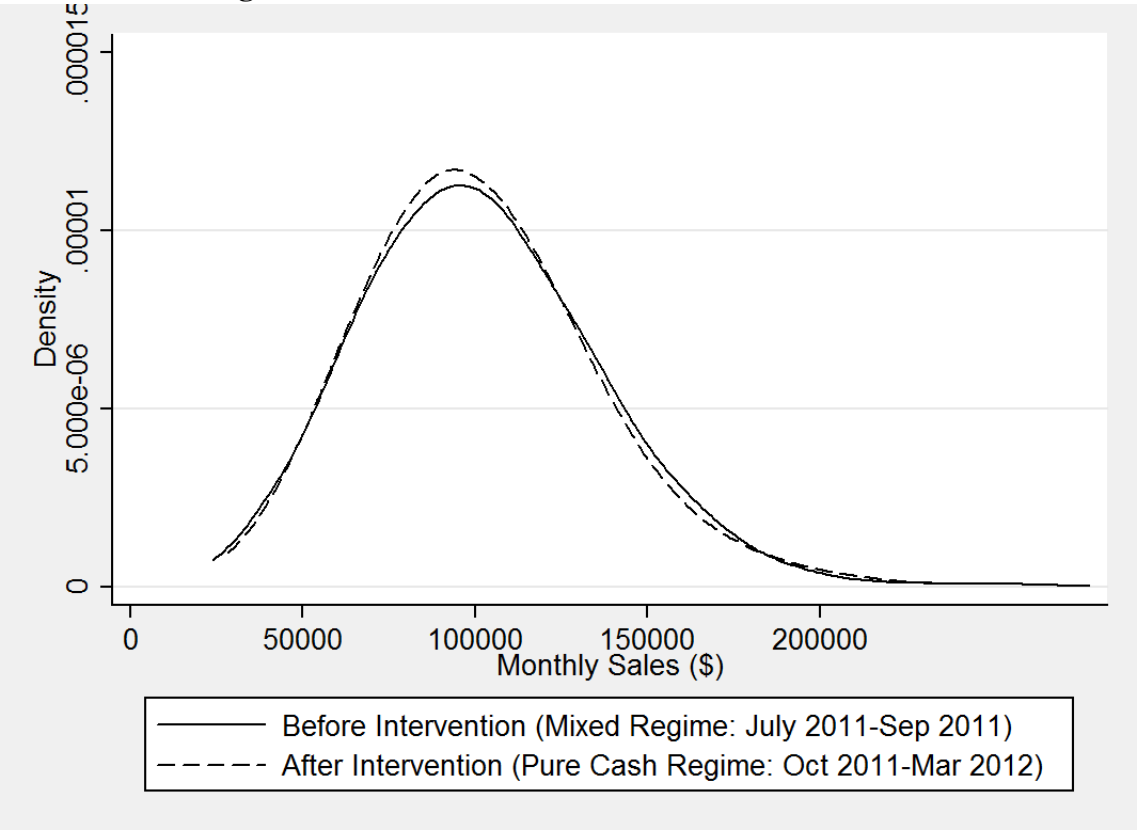


Figure 3: CATEGORY versus FIRM SALES

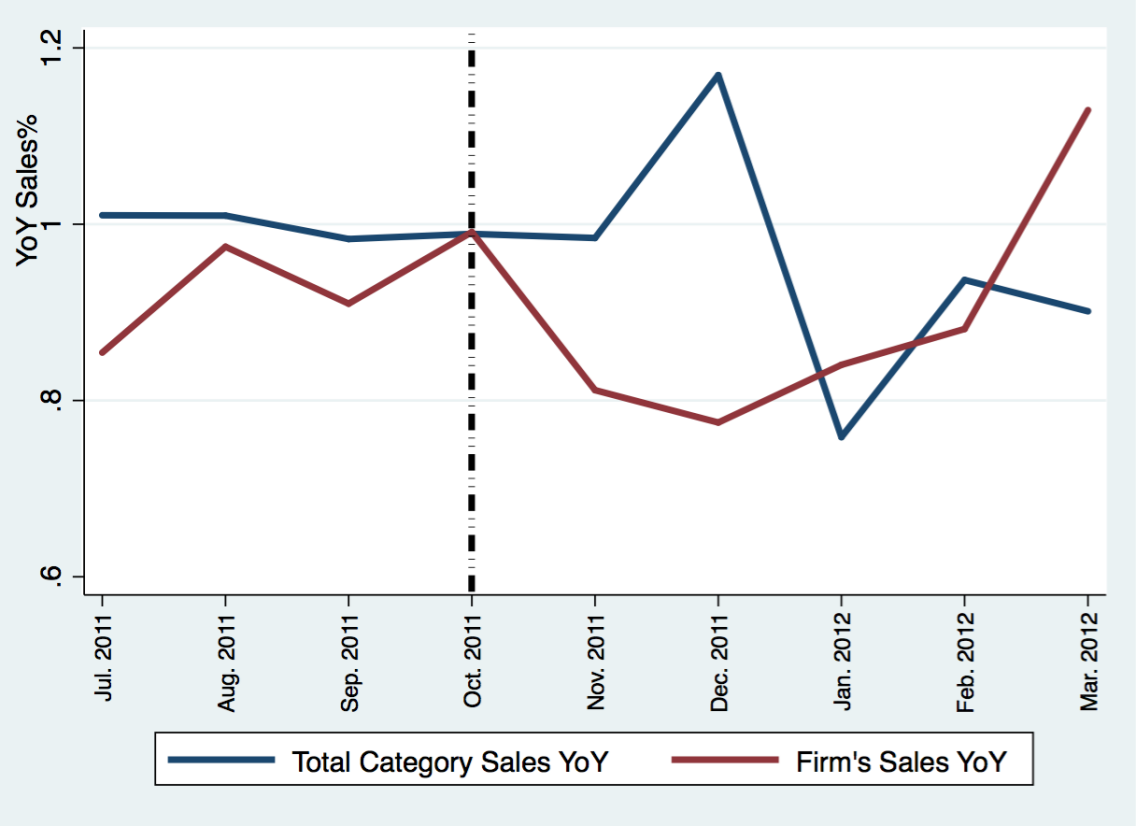


Figure 4: MARGINAL VALUE OF POINTS ACROSS REWARD LEVELS

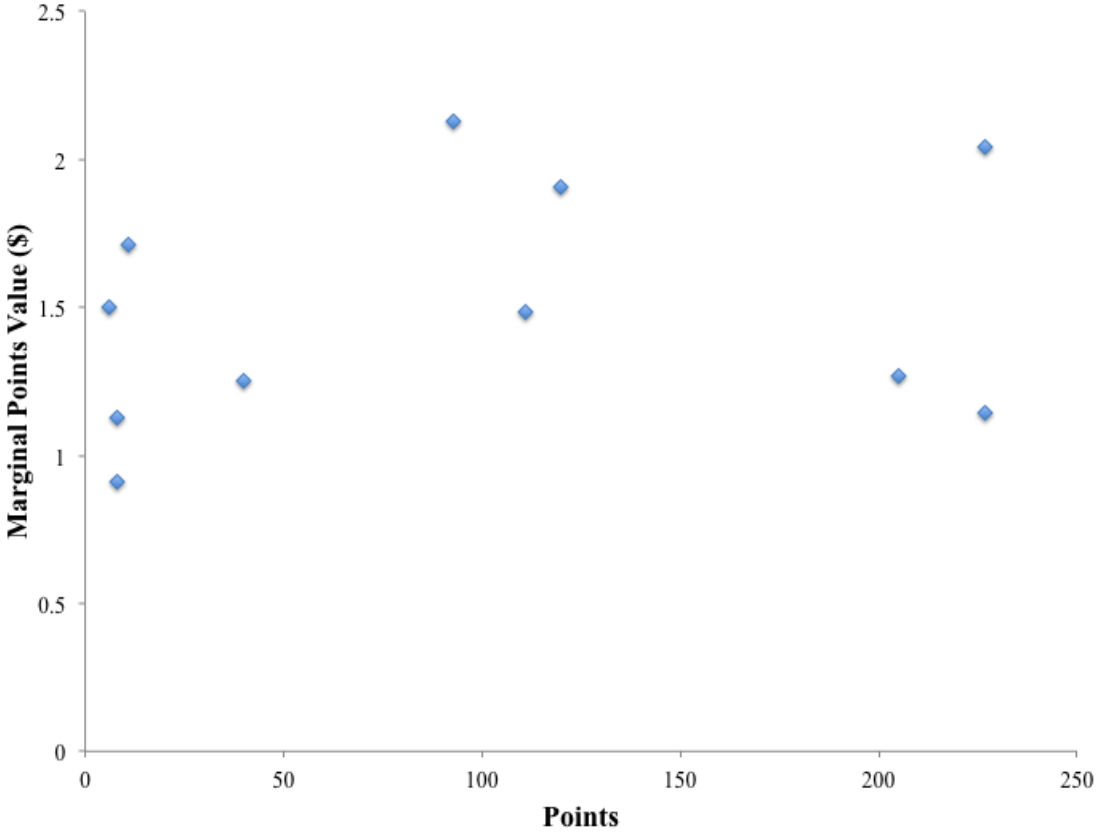


Figure 5: DISTRIBUTION OF INDIVIDUAL-LEVEL REGIME EFFECT

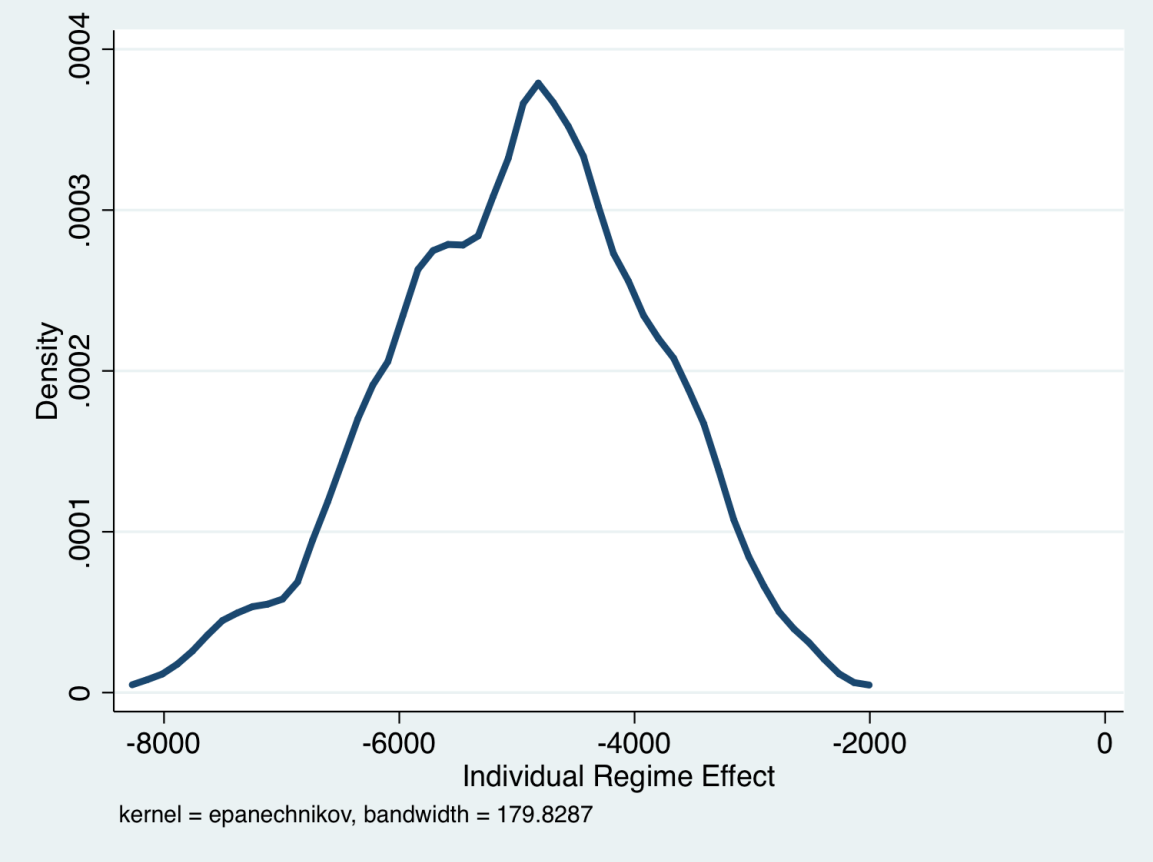


Figure 6: EFFORT DISTRIBUTION

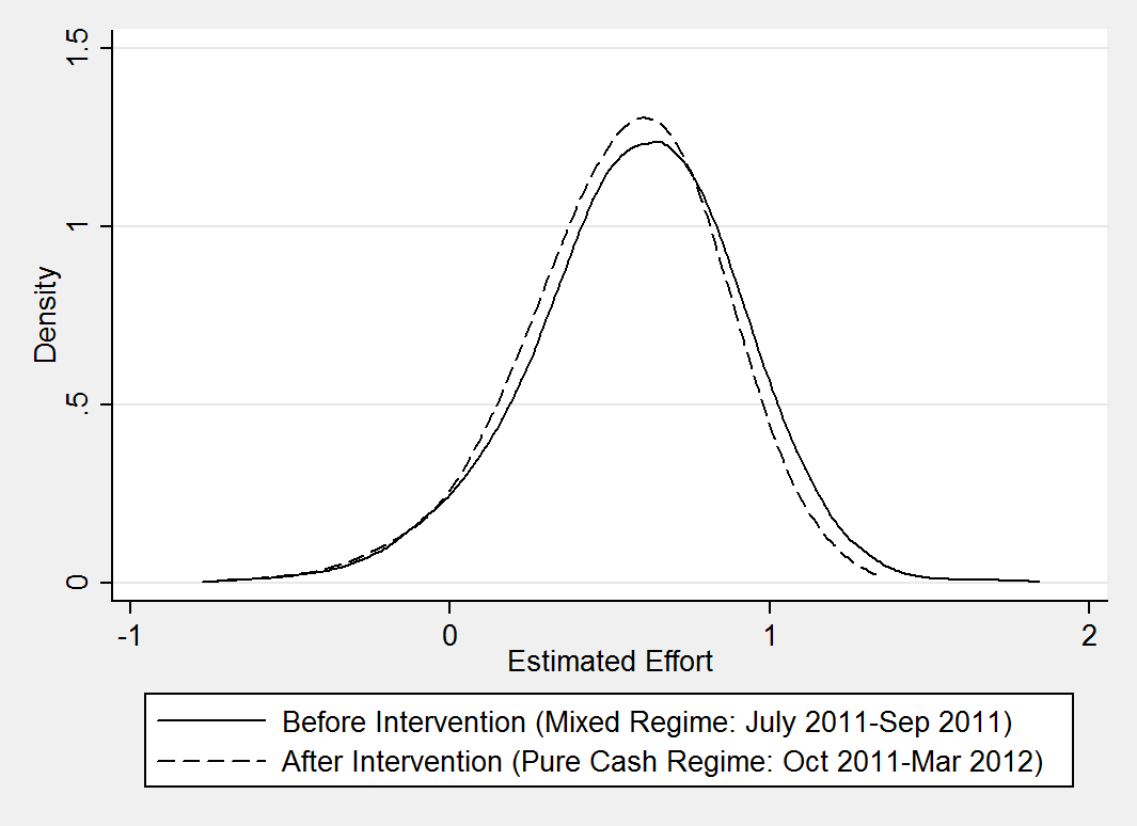
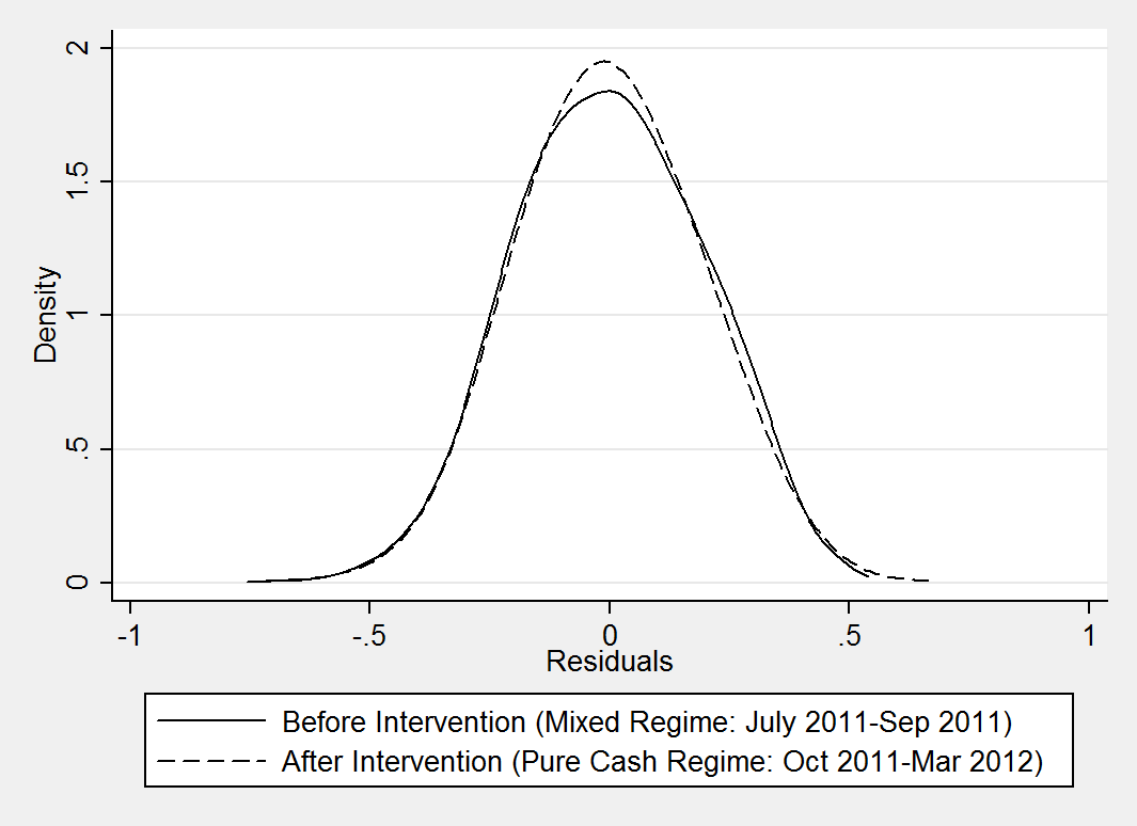


Figure 7: DISTRIBUTION OF RESIDUAL CHANGE



Web Appendix A: Robustness checks

Table WA.1: Robustness Check: Ratcheting Effect

Dependent Variable: Monthly Target (S)	
Independent Variable	Coefficient (Std. Err.)
Constant	-19,491.44 (6536.57)***
$Target_{t-1}$	1.12 (0.09)***
s_{t-1}	0.15 (0.12)
$Target^2_{t-1}$	-4.22E-06 (5.17e-07)***
s^2_{t-1}	-4.49E-07 (1.00e-06)
$Target^3_{t-1}$	7.45E-12 (9.68e-13)***
s^3_{t-1}	1.608E-12 (2.71e-12)
MonthSalesPriorYear	0.07 (0.01)***
Year Dummies	Included
Salespeople Fixed Effects	Included
Adj. R^2	0.88

Number of observations = 4,640

Table WA.2: Robustness Check: Reactivity

Dependent Variable: Monthly Sales ($)	
Independent Variable	Coefficient (Std.Err.)
Constant	-34,581.56 (15.519.52)*
Regime	-5,234.81 (867.55)***
MonthSalesPriorYear	0.03 (0.002)***
Target	-0.07 (0.01)***
Accumulated Points	325.99 (255.05)
Accumulated Points x Regime	0.25 (4.95)
Year Dummies	Included
Salespeople Fixed Effects	Included
Adj. R^2	0.66

N of Obs = 5,220

Table WA.3: Robustness Check: Sorting (Fixed Coefficients)

Dependent Variable: Monthly Sales ($)	
Independent Variable	Coefficient (Std. Err.)
Regime	-2,197.49 (633.75)***
Target	0.25 (0.01)***
MonthSalesPriorYear	-0.04 (0.01)***
Year Dummies	Included
Salesperson Fixed Effects	Included
Constant	38,833.39 (6,912.79***
Adj. R^2	0.65
N of Obs = 5,934; N of Routes = 676	

*** Significant at 99% level.

Table WA.4: Impact of Intervention on Sales

Dependent Variable: Monthly Sales ($)	
Independent Variable	Coefficient (Std. Err.)
Regime	-7,952.40 (653.68) ***
Target	0.20 (0.01) ***
MonthSalesPriorYear	-0.12 (0.01) ***
Price	115,438.89 (4490.73) ***
Constant	-356,460.09 (16965.85) ***
Year Dummies	Included
Salesperson Fixed Effects	Included
Adj. R^2	0.66
Number of observations	5220
# of salespeople=580	

*** Significant at 99% level.

Table WA.5: Impact of Intervention on Sales

Dependent Variable: log(Monthly Sales)	
Independent Variable	Coefficient (Std. Err.)
Regime	-0.04 (0.01) ***
log(Target)	0.16 (0.02) ***
log(MonthSalesPriorYear)	-0.05 (0.02) ***
Constant	9.62 (0.24) ***
Year Dummies	Included
Salesperson Fixed Effects	Included
Adj. R^2	0.70
Number of observations	5220
# of salespeople=580	

*** Significant at 99% level.

Web Appendix B

Table WB.1: Correlation Matrix

	Monthly Sales	Target	MonthSalesPriorYear	Firm-Level Price Index
Monthly Sales	1.0000			
Target	0.5713	1.0000		
MonthSalesPriorYear	0.5378	0.7194	1.0000	
Firm-Level Price Index	-0.0420	0.0709	0.0126	1.0000

Table WB.2: Intervention Effect (Unconditional)

Dependent Variable: Territory Monthly Sales ($)	Coefficients (Std. Err.)	
Independent Variable	Model 1	Model 2
Regime	-6,601.06(937.33)***	-5,451.41(1099.32)***
Target	No	No
$Target^2$	No	No
$Target^3$	No	No
MonthSalesPriorYear	No	No
Lagged Monthly Sales	No	No
Year Effects	No	Included
Salesperson Effects	No	No
Constant	101,686.60(782.44)***	101,686.60(777.33)***
Adj. R^2	0.01	0.01
N of Obs.	5,220	5,220
N of salespeople = 580	*** significant at 99%	

Table WB.3: Intervention Effect (Fixed Coefficients; Sample with Turnover)

Dependent Variable: Territory Monthly Sales ($)	
Independent Variable	Coefficient (Std. Err.)
Regime	-4,982.72 (611.37) ***
Target	0.3 (0.01) ***
MonthSalesPriorYear	-0.10 (0.01) ***
Firm-Level Price Index	106,953 (4,395.25) ***
Year Effects	Included
Salesperson Fixed Effects	Included
Constant	-335,633.78 (16,724.93) ***
Adj. R^2	0.64
Number of observations	5,934
Number of salespeople = 676	

*** Significant at 99% level.

Table WB.4: Intervention Effect (Random Coefficients, Sample with Turnover)

Dependent Variable: Territory Monthly Sales ($)		
	Independent Variable	Coefficient (Std. Err.)
Fixed Part		
	Regime	-5,334.94 (640.66) ***
	Target	0.048 (0.01) ***
	MonthSalesPriorYear	0.07 (0.01) ***
	FirmLevelPriceIndex	-104,004.64 (4,597.70) ***
	Salesperson Effects	329,538.69 (16,164.96) ***
Random Part		
	σ_β[Regime]	0.002 (0.19)
	σ_α[Salesperson Effect]	8,643.09 (698.62) ***
Pseudo R^2	0.55	
N of Obs = 5,934	N of salespeople = 676	*** significant at 99% level

Table WB.5: Impact Of Intervention On Sales (Random Coefficient Model)

Dependent Variable: Monthly Sales in Dollars($)		
	Independent Variable	Coefficient (Std. Err.)
Fixed Part		
	Regime	-7839.57 (671.45) ***
	Target	0.35 (0.01) ***
	MonthSalesPriorYear	-0.003 (0.01)
	Firm-Level Price Index	112493.66 (4610.07) ***
	Salesperson Effects	-334016.44 (16220.74) ***
	Year Effects	Included
Random Part		
	σ_β[Regime]	1.50E-05 (1.55E-05)
	σ_α[Salesperson Effect]	15015.81 (838.40) ***
Pseudo R^2		0.50
Number of observations=5220	(# of salespeople=580)	

*** Significant at 99% level.[Model 1: RC with only regime random] [Model 2: Everything random]

Robustness Check with log-log functional form

Table WB.6: Impact of Intervention on Sales (Random Coefficient Model)

Independent Variable		Dependent Variable: log(Monthly Sales) in Dollars Coefficient (Std. Err.)
FIXED PART		
	Regime	-0.04 (0.01) ***
	Target	0.49 (0.02) ***
	MonthSalesPriorYear	0.16 (0.02) ***
	Salesperson Effects	3.95 (0.16) ***
	Year Effects	Included
RANDOM PART		
	σ_β[Regime]	9.34E-09 (8.14E-09)
	σ_α[Salesperson Effect]	0.09 (0.01) ***
Pseudo R^2		0.52
Number of observations=5220	(# of salespeople=580)	

*** Significant at 99% level.[Model 1: RC with only regime random] [Model 2: Everything random, Std. err. cannot be calculated]

Figure 8: DISTRIBUTION OF FIRM MONTHLY SALES

Web Appendix C: Survey Questionnaire

The following questions were asked on the questionnaire.

Discretionary Expenditure: On average, what percent of your (family's) income is spent on discretionary expenditure every month? Discretionary expenditure refers to expenses on entertainment, restaurants, gifts, movies and electronics?

Household Size: How many people, including yourself are there in your household?

Redemption Frequency: Which statement best describes your Ovation redemption behavior?

- I typically redeem most of my Ovation points once a year
- I typically redeem most of my Ovation points once a quarter
- I typically redeem most of my Ovation points once a month
- I rarely/never redeem my Ovation points

Redeemed Merchandise: What were the last two items you redeemed using Ovation points?

We coded the responses as follows. The 546 items that were redeemed consisted of 230 unique items.To classify the items as hedonic or utilitarian, we hired independent researchers to categorize these items as either "utilitarian" or "hedonic". The instructions provided to the researchers were:

Hedonic goods provide more experiential consumption, fun, pleasure, and excitement, the ones whose consumption is primarily characterized by an affective and sensory experience of aesthetic or sensual pleasure, fantasy, and fun.(Something that is enjoyable, e.g., perfume, designer clothes, sports cars, luxury watches, etc.)

Utilitarian goods are primarily instrumental and functional, the ones whose consumption is more cognitively driven, and goal oriented and accomplishes a functional or practical task. (Something that is useful, practical, functional, that helps achieve a goal, e.g., a vacuum cleaner)

Please read the following list of products, which are redeemed products for salespersons as part of their bonus. Categorize each product as "1" (hedonic) or "2" (utilitarian).

Abstraction Level When you think about money, how likely are you to think about (7-point scale from "Dollar bills/coins" to "Purchasing goods and services")

Conjoint Questionairre The following questions were asked in the conjoint.

Your current compensation plan has 4 items; a) Commission on units sold, b) A cash bonus for target achievement, c) Cash payments for every percent increase in $ sales over target, and d) Ovation points for every percent increase in $ sales over target.

We would like to understand your preferences for items (c) and (d). For the plans below, tell us which one you would prefer by clicking the button below it.

In each plan, assume all unspecified items are maintained at same levels for all choices.

1. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $10 PLUS 20 Ovation Points - PLAN A

   (b) $0 PLUS 30 Ovation Points - PLAN B

   (c) $0 PLUS 20 Ovation Points - PLAN C

2. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $10 PLUS 30 Ovation Points - PLAN A

   (b) $0 PLUS 0 Ovation Points - PLAN B

   (c) $30 PLUS 20 Ovation Points - PLAN C

3. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $30 PLUS 0 Ovation Points - PLAN A

(b) $20 PLUS 10 Ovation Points - PLAN B

(c) $10 PLUS 10 Ovation Points - PLAN C

4. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $10 PLUS 30 Ovation Points - PLAN A

   (b) $0 PLUS 0 Ovation Points - PLAN B

   (c) $30 PLUS 20 Ovation Points - PLAN C

5. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $0 PLUS 10 Ovation Points - PLAN A

   (b) $0 PLUS 20 Ovation Points - PLAN B

   (c) $10 PLUS 0 Ovation Points - PLAN C

6. The following options show the $ and ovation points you would get for every percent increase over $ sales. If these were your only options, which would you choose? Choose by clicking one of the buttons below:

   (a) $10 PLUS 10 Ovation Points - PLAN A

   (b) $20 PLUS 0 Ovation Points - PLAN B

   (c) $10 PLUS 20 Ovation Points - PLAN C

Web Appendix D

Estimating Effort Estimates

Monthly sales in month t under regime $\{s \in mixed, pure\}$ generated by each salesperson is :

$$log(Sales_{it}) = \alpha_i + \beta_2 Regime_t + \beta_3 Year_t + \beta_4 log(MonthSalesPriorYear_{it}) + \beta_5 log(target_{it}) + \beta_6 [Regime_t \cdot \alpha_i] + \varepsilon_{it}$$

where

$$Regime_t = \begin{cases} 0 & \text{mixed regime} \\ 1 & \text{pure cash regime} \end{cases}$$

We assume that salespersons' effort:

$$\hat{e}_{it}^s = \hat{\alpha}_i^s + \hat{\varepsilon}_{it}^s$$

where

$$\hat{\alpha}_i^s = Regime_t * \hat{\alpha}_i$$

$$\hat{\varepsilon}_{it}^s = Sales_{it} - \hat{Sales}_{it}$$

Decomposing effort change

If we look at fixed effects and residual separately, the residual distribution is centered around 0, but variance dropped (standard deviation before = 0.2756, standard deviation after = 0.2546). Therefore, the drop on effort comes from fixed effect change.

Calculating the average effort change Average effort is given by:

$$\hat{e}_{it} = \sum_{t=1...T} \hat{e}_{it}$$

Among 580 agents, we can identify the effort change for 579 agent, among which, 370 (63.90%) salespeople's effort dropped (on average, effort decreased by -10.70%), and only 209 (36.10%) salespeople's effort increased (on average, effort increased by 7.45%). (t-test p=0.0000)

Table WD.1: EFFORT ESTIMATES

	MEAN	MEDIAN	STD. DEV.
Average effort Before Intervention ($\hat{e}_{iMixed}$)	0.58	0.60	0.32
Average effort After Intervention ($\hat{e}_{iPureCash}$))	0.54	0.56	0.31
Effort Change ($\hat{e}_{iPureCash} - \hat{e}_{iMixed}$)	-0.04	-0.04	-0.02
N obs	579		

T-test for comparing effort changes

$$\begin{aligned} t &= \frac{\overline{\hat{e}_{after}} - \overline{\hat{e}_{before}}}{\overline{Var[\hat{e}_{after} - \hat{e}_{before}]}} \\ where & \\ Var[\overline{\hat{e}_{after}} - \overline{\hat{e}_{before}}] &= \frac{Var[\hat{e}_{after} - \hat{e}_{before}]}{N^2} \\ &= \frac{Var[\hat{e}_{after}] + Var[\hat{e}_{before}] - 2Cov[\hat{e}_{after}, \hat{e}_{before}]}{N^2} \end{aligned}$$

Assuming $Cov[\hat{e}_{after}, \hat{e}_{before}] = 0$,

$$Var[\overline{\hat{e}_{after}} - \overline{\hat{e}_{before}}] = \frac{Var[\hat{e}_{after}] + Var[\hat{e}_{before}]}{N^2}$$

Given that $se(\hat{e}_{iafter}) = s_{iafter}$ and $se(\hat{e}_{ibefore}) = s_{ibefore}$, we calculate the t-test for the following hypothesis.

$$\begin{aligned} H_0 &: e_{after} - e_{before} = 0 \\ H_1 &: e_{after} - e_{before} \neq 0 \end{aligned}$$

The results of the t-test are $t = -8.37, p = 0.000$ indicating that effort expended before the intervention was significantly more than the effort expended after the intervention.