Institute for Policy Research
Northwestern University
Working Paper Series

WP-14-06

Reducing Moral Hazard in Employment Relationships: Experimental Evidence on Managerial Control and Performance Pay?

Kirabo Jackson
Associate Professor of Human Development and Social Policy
Faculty Fellow, Institute for Policy Research
Northwestern University

Henry Schneider
Assistant Professor of Economics
Cornell University

Version: November 2013

DRAFT
Please do not quote or distribute without permission.

Abstract

Moral hazard is endemic to employment relationships and firms often use performance pay and managerial control to address this problem. While performance pay has received much empirical attention, managerial control has not. Jackson and Schneider analyze data from a managerial-control field experiment in which an auto-repair firm provided detailed checklists to mechanics and monitored their use. Revenue was 20 percent higher under the experiment. They compare this effect to that of quasi-experimental increases in mechanic commission rates. The managerial-control effect is equivalent to that of a 10 percent commission increase. They find evidence of complementarities between the two, suggesting benefits from an all-of-the-above approach. They also find evidence of incentive gaming under performance pay.

Economists have long recognized that employment relationships are beset by moral hazard. When worker and firm incentives are not aligned, and worker effort is imperfectly observed, workers may exert inefficiently low levels of effort. Tying worker compensation directly to firm outcomes via performance pay can help align these incentives (Mirrlees 1999). Nearly all Fortune 1000 companies now use performance pay (Web and Blandin 2005) and popular acceptance of its benefits motivates current proposals to tie teacher pay to student test scores (Jackson, Rockoff, and Staiger 2014) and hospital compensation to patient outcomes (Mulen, Frank, and Rosenthal 2009).[2] Nevertheless, performance pay generally does not allow for the Pareto-optimal outcome when worker effort is imperfectly observed (Holmstrom 1979).

Because the root of the problem is the inability of firms to observe and hence effectively manage worker behavior (Arrow 1985), an alternative approach to reducing moral hazard is to increase worker monitoring and supervision. In models of worker shirking (e.g., Yellen 1984), monitoring helps firms detect shirking, thus incentivizing effort. In agency models (e.g., Stiglitz 1975, Harris and Raviv 1978, Holmstrom 1979) monitoring enables contracting on a more complete set of worker actions. These models often assume the workers' tasks are clearly defined and known to the worker and firm. In real-world settings, however, these tasks may be less explicit. As such, we define *pure monitoring* as observing worker actions to ensure compliance with prescribed tasks, and we define *managerial control* as the provision of explicit instructions to workers coupled with monitoring. In simple settings where the tasks are common information (as in many theoretical models), monitoring and managerial control are the same.

Despite theoretical findings that performance pay and managerial control may mitigate moral hazard, empirical research has focused almost exclusively on performance pay (e.g.,

[2] This logic also motivates the use of stock options for CEOs (Mishel and Sabadish 2012), and the use of commission pay for sales jobs (Eisenhardt 1988) and other occupations (Lemieux, MacLeod, and Parent 2009).

Foster and Rosenzweig 1994, Lazear 2000, Shearer 2004).[3] Research related to managerial control includes Hubbard (2000) who found that demand for monitoring technologies was higher when workers had lower-powered incentives, and Bloom et. al. (2013) and Bruhn, Karlan, and Schoar (2013) who found that consulting services, which include bundles of business management practices such as monitoring and performance pay, increased firm productivity. Nevertheless, there is little direct evidence on the effects of increased managerial control on worker productivity, the relative merits of managerial control and performance pay, and any complementarities that may exist. This paper aims to fill this gap.

We analyze data from a field experiment conducted at a chain of auto-repair shops in which the firm increased managerial control of workers.[4] We directly compare the effects of managerial control and performance pay by exploiting quasi-experimental variation in commission rates for the same mechanics. We examine data on customers, cars, repairs, and mechanics for each car visit to 11 auto-repair shops over a five-year period. Our main outcome is revenue generated by the mechanic. We also examine measures of mechanic effort, the allocation of effort across tasks, and take-home-pay. Auto repair is an attractive industry for studying moral hazard because (1) mechanic actions are not precisely observable to the firm and so moral hazard is likely to be present;[5] (2) mechanic output is observable through outcomes such as repairs and revenue; and (3) the use of performance pay is widespread.

In the managerial-control experiment, three shops were randomly chosen to receive treatments. Treatment mechanics were instructed to use checklists for approximately one-month, which would be collected by the supervisor. The checklists contained a list of car components

[3] The literature also includes a number of field studies on performance-pay-related topics including Bandiera, Barankay, and Rasul (2005, 2008), Lindenauer et. al. (2007), and Hossain and List (2012).

[4] In related work, Schneider (2010) and Johnson, Schneider, and Waldman (2013) finds considerable moral hazard in leasing markets, while Jackson and Schneider (2011) find that social pressure can mitigate this moral hazard.

[5] Schneider (2012) finds that misdiagnosis is a fundamental characteristic of the auto-repair market.

and were designed to guide mechanics through a thorough car inspection. The checklists made explicit to the mechanics the steps the firm wished them to conduct, and were designed to ensure that mechanics were exerting sufficient effort to identify car problems that generate repair work. Collecting the completed checklists allowed the firm to monitor mechanic compliance.

While checklists can be memory aids and have been found to reduce errors in medicine and aviation (e.g., Boorman 2001, Gawande 2009), key aspects of successful checklist interventions are to make explicit the tasks that must be completed and to facilitate greater monitoring and accountability.[6] Accordingly, the managerial-control treatment reflects the effects of (1) instructing workers on which actions to take (the steps of the inspection); (2) instructing workers to use a memory aid to adhere to these instructions; and (3) monitoring worker compliance with the instructions. Note that if workers were fully aware of how to conduct a thorough inspection and aware of the benefits of checklists, explicit instructions would not be necessary, and the managerial-control treatment would be a pure monitoring intervention.

We conduct both an event study and a difference-in-differences analysis, comparing the change in outcomes before, during, and after the treatment to the change in outcomes for untreated workers. During the treatment, checklist use increased from zero to 33 percent of cars, and revenue generated by treated mechanics rose 20.2 percent ($p<0.01$). After the treatment, checklist use returned to zero and revenue reverted to pre-intervention levels. We present tests to show robustness to small-sample inference and to support a causal interpretation.

To measure the effects of an increase in performance pay, we examine instances in which mechanics received commission increases. We use data on the exact dates of the commission increases for each mechanic in combination with the longitudinal nature of the data. This

[6] For example, in the Michigan Keystone ICU program, nurses were empowered to intervene if doctors were not following the steps of a checklist (Pronovost et. al. 2006, Bosk et. al. 2009).

facilitates both an event-study analysis and a difference-in-differences analysis where we compare the change in a mechanic's outcomes after a commission increase to the change in outcomes for other mechanics that did not receive a commission increase at that time. These within-worker approaches allow us to attribute observed changes in productivity to changes in mechanic behavior. Because the commission increases did not occur randomly, we present several tests to establish that the results can be interpreted causally.

A one percentage point increase in the commission rate (a six percent increase relative to the base commission rate) increased revenue by 11.7 percent ($p<0.01$). As such, the managerial-control effect is equivalent to the effect of a 10 percent increase in the commission rate. In a simple model with a single task (e.g., mechanics only perform repairs), managerial control and performance pay may be substitutes because both increase effort for the same task. However, with multiple complementary tasks (e.g., mechanics both inspect and repair the car), managerial control may increase effort on inspections while performance pay may increase effort on repairs such that the two are complements. Testing for an interaction effect, we find that the managerial-control effect was larger for mechanics with a higher commission rate, suggesting that managerial control and performance pay are complements in this context.[7]

We next investigate mechanisms. The managerial-control treatment led to an increase in worker effort as measured by the numbers of repairs conducted and hours worked, while commission increases led to a shifting of effort toward repairs that generated more revenue, without associated increases in number of repairs or hours worked. Insofar as this shifting toward more expensive repairs reflects mechanics exploiting their informational advantage over

[7] From an organizational perspective, Baker and Hubbard (2004) examined the trucking industry and found that truck ownership decreased as the use of on-board computing grew. Similarly, Krueger (1991) found that workers for whom monitoring was less complete received higher compensation rates and steeper tenure-earnings profile. Both results suggest a substitute nature of monitoring and monetary incentives.

customers (Hubbard 1998, Schneider 2012), the results underscore the possibility of deleterious side effects of performance pay.[8]

The results indicate the presence of moral hazard in the firm-employee relationship. Moral hazard arises because mechanics receive only a share of the revenue they generate but incur the time and effort costs of their work. The plainest evidence of this moral hazard is that mechanics chose to discontinue checklist use and revenue fell back to pre-treatment levels after the treatment periods ended, despite the large revenue gains for the firm from the treatment. We also calculate, based on income received and hours worked by mechanics, that the thorough inspections benefit the firm but not the mechanics under the current compensation structure.

This study contributes to our understanding of moral hazard and employment relations. First, it is the first to demonstrate directly the ability of increased managerial control to reduce moral hazard. Second, it contributes to the literature on management practices (e.g., Lucas 1978, Bertrand and Schoar 2003, Bloom et. al. 2013) by identifying the effect of specific procedures that are often components of bundled management interventions. Third, the findings contribute to the literature on performance pay by providing new evidence of its efficacy and potential side effects. Finally, it presents direct field evidence of the theoretical result that the optimal incentive contract will depend on the level of managerial control.

The paper proceeds as follows. Section I describes the data and the industry. Section II describes the experiments. Sections III and IV present the main results and robustness checks. Section V examines mechanisms. Section VI concludes.

[8] Another motivation for studying alternate approaches to reducing moral hazard is that performance pay can induce undesirable worker behavior when the performance being rewarded does not capture all dimensions of firm output (Holmstrom and Milgrom 1991, Baker 1992). For example, teachers that are rewarded based on student test scores might teach to the test instead of promoting deeper learning (Glewwe, Ilias, Kremer 2010), and doctors rewarded for better patient outcomes may avoid high-risk patients (Rosenthal and Frank 2006). Other empirical papers on the trade-offs of performance pay include Jacob and Levitt (2001), Freeman and Kleiner (2005), Wright et. al. (2007), Mullen, Frank, and Rosenthal (2010), and Larkin (forthcoming).

I. DESCRIPTION OF THE AUTO-REPAIR BUSINESS AND DATA

We worked with 11 shops in one metropolitan area from a United States auto-repair chain. The shops are centrally owned. The firm provided data on all customers, cars, repairs, and employees for the period June 23, 2008 and June 22, 2013. This includes information about every repair, part, and charge; the mechanic and manager and associated pay structure, daily minutes worked, and weekly compensation amount; car characteristics including make, vintage, and mileage; and customer characteristics such as first name, from which we derive gender.

The source of revenue for the repair shops is repair charges.[9] Table 1 shows the repair categories that had repairs in at least one percent of visits, and the mean charge for each. Roughly half of visits involve oil changes, followed by brake repairs (15 percent) and alignment/suspension repairs (12 percent). Among visits with oil changes, 37 percent involve additional repairs. A primary aim of the firm is to convert routine maintenance visits (e.g., oil changes) to include additional repairs that are discovered while the car is in the shop.

A visit is defined as an invoice, which is essentially one contiguous repair visit by a particular car. An invoice may be a simple oil-change visit or can involve the customer leaving the car overnight or returning several days later for scheduled work. The mean amount charged per visit, or mean revenue per visit, is $190.61. This amount is labor charges plus parts charges minus any coupons or other discounts, where 52 percent of visits involve some discount. Thus, the revenue amount is what the customer actually pays. The results are similar when discounts are excluded. Repairs range from the very inexpensive, such as oil changes and windshield

[9] The firm uses a rule-of-thumb to set repair charges such that the internal cost of parts represent approximately 20 percent of the total repair charge. The remaining 80 percent encompasses the labor charges based on the shop's posted hourly labor rate and the standard book time for that repair, and a mark-up that is allocated to parts charges. For example, if the internal cost to the shop for a part is $10, the standard labor time for that repair is half an hour, and the posted hourly labor rate at the shop is $70, then the customer would be charged around $90 total, with $35 reported to the customer as labor charges and $55 reported as parts charges.

wipers, which have mean charges of $24 and $25, respectively, to labor and parts-intensive work such as brake and suspension repairs, which are $300 or more.[10]

There is a single owner for all shops, and a shop manager for each shop who supervises the mechanics and interacts with customers. Each shop has two to three mechanics on duty in the shop at any time who conduct the inspections and repairs. We focus on the 108 mechanics who worked at the shops during the sample period.[11] Table 2 presents summary statistics on all mechanics during the sample period. Seventy percent of visits are handled by mechanics who are paid primarily on commission. Mechanics who work primarily on commission ("commission mechanics") receive the maximum of a commission and an hourly rate, such that their entire pay is commission except in rare particularly slow weeks where the hourly rate ensures a minimum pay amount. These mechanics have a commission rate between 14 to 20 percent, with a mean of 17.6 percent, and an hourly rate between $9 to $12.[12,13] Most of the remaining visits are split between mechanics that are paid on an hourly rate ($9 to $12 per hour) or a flat rate ($20-$24 per hour of labor billed to customers). On average, a mechanic works 47.2 hours per week over 4.90 days, and has a weekly pay of $798. The mean number of visits per mechanic at the firm since 1998 is 3,128, and average mechanic tenure is 3.7 years (including mechanics still working at the firm). Commission mechanics have even more experience, having worked on 5,038 cars and for 5.2 years at the firm on average.[14]

---

[10] Appendix Table A1 reports the most common invoice items in the sample. Figure A1 shows revenue per visit and number of visits per shop by month for 2009-2011. A seasonal trend appears, which will be important to control for.
[11] These 108 mechanics includes 25 managers in our sample who only occasionally conducted repairs, and who are not included in Table 2 to avoid distorting several statistics in the table (e.g., number of visits per day).
[12] Mechanics are paid every Wednesday for work over the prior Monday to Saturday. Shops are closed on Sundays.
[13] Two mechanics whose compensation is primarily hourly also receive commissions, of 2 to 3 percent. Managers, who occasionally conduct repairs, have a commission rate for those repairs of 10 percent.
[14] Therefore, turnover is less central for this firm than some others such as in Lazear (2000). Regardless, we examine worker-level outcomes, and so changes in worker composition do not directly play a role in our estimation.

While most of the data are complete, the payroll data set, which reports time worked and weekly pay, contains data only on mechanics that are currently employed by the firm, and so this part of our analysis is based on approximately 60 percent of mechanic-weeks over the five-year period, skewed toward to the more recent period. Our results for the revenue, repair, and other analyses are very similar when those data are limited to mechanics still employed by the firm.[15]

II. THE EXPERIMENTS

A primary objective of this study is to estimate the causal effects of an increase in managerial control and an increase in incentive pay for the same population of workers and set of outcomes. The following two experiments provide independent sources of variation for estimating these effects: (1) a controlled experiment in which managerial control was increased, and (2) a quasi-experiment in which commissions were increased.

A. The managerial-control experiment

To identify the effects of increasing managerial control we worked with the firm owner to implement a series of checklist interventions at three of the 11 shops. The owner asked the mechanics to use and fill-out checklists for cars that came into the shop. The mechanics were told to use the forms on as many cars as they comfortably could, but were not told to use the forms on every car, nor were mechanics told to use checklist on any particular type of car. An example of a checklist visit is as follows: a customer brings in her car for an oil change, and while the customer is waiting, the mechanic inspects the car and completes the checklist. The mechanic then shows the completed checklist to the manager and the manager suggests the

[15] Note that we worked with the same firm on a different experiment that occurred primarily in 2013. For completeness, in all of specifications we include a dummy variable for the second experiment so as to difference out any of these other treatment effects. Results are unchanged when these treatment periods are excluded.

repairs to the customer. The customer consents to none, some, or all of the recommended repairs. There were 655 visits during the treatment periods, of which 214 or 32.7 percent had a checklist. Given the possible selection of which cars received checklists during the treatment period, we consider all visits during the treatment period as treatment visits, and compare these outcomes to those in the non-treatment period.

The rationale for the checklists is to induce the mechanics to conduct more thorough inspections in order to identify more potential repair work.[16] Many other auto-repair firms use checklists, and the supervisor told us that all mechanics were well aware of checklists and none at the shops in our sample used them.[17] Knowing that simply providing checklists would not generate meaningful checklist use, the owner required mechanics to submit the completed checklists to a supervisor every week, which were subsequently returned to the firm headquarters.[18] This is the monitoring aspect of the intervention. The subjects were aware they part of an experiment run by the firm with assistance from academics to learn about checklist efficacy and so the intervention could be considered a framed field experiment.[19]

There were four managerial-control treatments at three shops (one shop had two treatments). Each treatment occurred for a series of days so that the treatment turned on and off, allowing us to observe the effect of both occurrences.[20] As Table 3 shows, these three shops were similar to the eight non-treatment shops. The general windows of treatment dates correspond roughly to when we were in contact with the firm. The exact treatment dates within the general

---

[16] In medicine, checklist use has been found to improve outcomes, it has been suggested by addressing issues of bounded rationality such as inattention, forgetting, or complexity (Gawande 2009) and increasing worker effort through greater monitoring, and collaboration (e.g., Bosk et. al. 2009).

[17] For example one can find auto-repair checklists with a simple Internet search. A sample link is below http://www.i-car.com/pdf/advantage/online/2003/042103_Checklist.pdf [last accessed on November 3, 2013].

[18] Copies of the completed checklist (with links to the associated invoice) were provided to us.

[19] A framed field experiment is defined as having suitable subjects in their natural environment with real incentives that are aware of the experiment (see Table 1 in List and Rasul 2011 for a taxonomy).

[20] Treatments occurred at shop 1 from July 10 to August 11, 2012, at shop 2 from March 3 to March 26, 2012 and November 5 to November 15, 2012, and at shop 13 from July 10 to August 6, 2012.

windows correspond to when the operations supervisor visited the shops for unrelated reasons, and was able to drop off and pick up the checklists at the same time.[21]

During the treatment, mechanics were given checklists that listed different car parts and maintenance items that the mechanics should inspect. The items include checking the oil level, various components of the brakes, wheel alignment, and so on. Two distinct checklist forms were used, which had very similar sets of items but different layouts and orderings. Checklist forms 1 and 2 were used in 34 and 66 percent of visits, respectively. Figures 5 and 6 show samples of checklist forms 1 and 2, respectively.[22] The high level of detail written on many of the completed checklists indicates that the mechanics actually used the checklists and did not simply fill them out with erroneous information. We also show in Section V that the checklist layout affected which repairs were conducted, again indicating that mechanics actually used them. While filling out the checklist itself is not time consuming (taking several minutes or less), actually conducting a more thorough inspection requires a non-trivial time cost for the mechanic.

In our primary analysis in Section III, we provide results from a range of regression models and outcomes. However, because the number of treatments is small, we first establish visually with an event-study analysis that the observed treatment effect is not driven by random chance. We regress the log of each invoice amount on mechanic and shop indicator variables to remove mechanic and shop-level means, and indicator variables for each year-week to remove any shocks that are common across shops. We then compute the mean residuals for the treatment mechanics during the treatment periods and for each month (four-week period) before and after

[21] The operations supervisor visits all shops every week or two for routine business operations.

[22] The checklists are similar to a typical state safety inspection form. However, the shops are in a state without mandated safety inspections and so many of the cars may not have received a thorough inspection in the recent past. Nineteen U.S. states currently require annual or biennial safety inspections for passenger cars, typically including a checklist for the inspection process. See http://en.wikipedia.org/wiki/Vehicle_inspection_in_the_United_States for more information (last accessed on 4/29/13).

the treatment periods. The mean residuals are re-centered to have a mean of zero in the ten non-treatment-period months. We chose a monthly period because this is approximately the mean treatment-period duration.[23] Thus, this residual plot depicts the *change* in mean invoice amount (net of shop, mechanic, and time effects) for mechanics in treatment shops relative to the *change* over the same period for mechanics in control shops.

Using the mean residual values for the 42 complete monthly periods that precede the earliest treatment period, we derive the 95-percent confidence interval of the random variability in monthly mean residuals. This confidence interval represents the range within which 95 percent of the monthly mean residuals would lie by random chance if the treatment had no effect.

Before showing the effect on revenue, we show the effect on the fraction of visits in which mechanics used a checklist. This is in the top panel of Figure 1. Consistent with increased checklist use due to the increased monitoring (i.e., checklist collection), checklist use jumped during the treatment period among treatment mechanics. Checklist use then returned to zero after the treatment period, showing that mechanics took no independent initiative to use checklists. If the managerial-control treatment had a causal effect on revenue, there should be a spike in revenue coinciding with the treatment period.

The bottom panel of Figure 1 shows the event-study graph for revenue per visit. There is a clear increase in revenue during the treatment period and a return to pre-treatment levels afterward. The underlying variability in the data suggests that this spike did not happen by random chance. The standard deviation of the monthly mean residual is 0.073 such that 95 percent of period means would lie within 0.15 and -0.15 by random chance. The actual deviation

[23] We exclude the second treatment period for shop 2 from the event-study analysis because the graphs use the periods before and after the treatment periods, causing the months around the first and second treatment periods for shop 2 to overlap. The second treatment period at shop 2 also only lasts for eleven days, compared to an average of 28 days for the other three treatment periods, which interferes with statistical inference.

from the non-treatment mean is 0.19, 2.6 times the standard deviation, and would happen by random chance with less than one percent probability. Indeed the mean residual for the treatment period is much larger than for any of the other months in the five-year sample period (the second-largest mean residual is 0.07).

The timing of this outlier spike in revenue at exactly the treatment period is evidence for the causal interpretation of the managerial-control treatment on revenue. Furthermore, because differential trending between treatment and control workers appears to be absent, the causal effect of this experiment should be identifiable within a difference-in-difference framework. We describe this empirical strategy in Section III. We also show in the top panel of Figure A3 in the Appendix the kernel density plots of revenue shifting for the managerial-control experiment.

B. The commission-increase quasi-experiment

To identify the effects of paying higher commissions we exploit within-worker variation in commission rates. For mechanics that received a commission increase during the sample period ("CI mechanics"), we compare their outcomes before and after the commission increase. This within-worker approach avoids using variation in commission rates across mechanics, who may have unobserved productivity differences. To account for any unrelated events that may have coincided with the timing of the commission increases, we compare the change in outcomes for CI mechanics to the change in outcomes for mechanics whose commission rate did not change over the same time period ("non-CI mechanics"). This approach is valid as long as the CI mechanics did not experience other changes at the same time as the commission increases that were not experienced by the non-CI mechanics. We provide evidence of this now.

We observe 17 commission increases among the 108 mechanics in our sample, which are shown in Table 4. The commission increases (with the exception of one) occur in single percentage point steps. Because the mean commission rate among commission mechanics is 17.6 percent, these one percentage point increases represent a compensation increase of approximately six percent. Interpreting any change in outcomes corresponding to the commission increases as causal requires that (1) the outcomes of CI mechanics were not already on an upward trajectory before the commission increases, and (2) differences in outcomes between CI and non-CI mechanics coincides with the timing of the commission increases. For evidence on these two conditions, we present a graphical event-study analysis.

For our event-study analysis, we regress the log of each invoice amount on mechanic and shop indicator variables to remove shop and mechanic-level means, and indicator variables for each year-week to remove common time variation. We then compute the mean of the residuals for each of the months (four-week periods) leading up to and following the commission increases for the CI mechanics. Figure 2 is a plot of these monthly mean residuals. The mean residuals are re-centered to have a mean of zero in the pre-increase period. Thus, this residual plot depicts the change in the mean invoice amount (net of shop, mechanic, and time effects) for the CI mechanics in response to the commission increases relative to the change for the non-CI mechanics for the same time period. We also compute the standard deviation of the monthly mean residuals for the 12 months prior to the commission increases, and plot the associated 95-percent confidence intervals.[24] That is, we show the range of values within which 95 percent of the monthly means will fall by random chance if the difference in outcomes between the CI and non-CI mechanics after a commission increase is zero.

The standard deviation of the pre-commission-increase monthly mean is 0.074 such that 95 percent of monthly means should be between -0.148 and 0.148 of the pre-commission-increase mean if the commission effect were zero. Consistent with no pre-existing trending, none

[24] We do not use more than 12 prior months because some increases occurred near the start of the sample period.

of the monthly means leading up to the commission increase are outside of this range. However, after the commission increase, there is a visible increase in the invoice amount for CI mechanics relative to non-CI mechanics. Four of the five post-commission-increase months are above the 95 percent level and the lowest of the post-commission-increase residual means is above the highest of the pre-commission-increase residual means. The bottom panel of Figure A3 in the Appendix shows the kernel density plots of this shifting in revenue for the commission-increase quasi-experiment.

In summary, Figure 2 provides evidence that (1) CI mechanics had larger increases in invoice amount after a commission increase relative to non-CI mechanics, (2) this increase is not due to a pre-existing trend difference between CI and non-CI mechanics, (3) the timings of the revenue increases and commission increases coincide exactly, and (4) the increase is sufficiently large and persistent that random chance is unlikely to explain it.

The internal validity of the commission analysis would be compromised if the commission increases coincided with other changes for the CI mechanics. This is unlikely because mechanics received commission increases as a reward for good work and to increase retention, and hence the increases are based on past rather than projected future performance. In any case, we report evidence in Section IV that commission increases were unrelated to other changes for CI mechanics. The event-study results indicate that the commission experiment is amenable to a difference-in-difference estimation strategy. We present this below.

III. MAIN RESULTS

A. Empirical strategy

We combine experimental variation in the level of managerial control with quasi-experimental variation in commission rates to measure the effectiveness of managerial control and performance pay for mitigating moral hazard. The key variation in the level of managerial control and the commission rate that we use is within a mechanic over time, and we employ a difference-in-difference regression framework. We estimate the change in outcomes at the time of the exogenous changes in managerial control and commission rate for the affected mechanics and compare this to the change in outcomes at the same times for the control mechanics. This comparison is captured in following equation, which we estimate by OLS,

[1] $$Y_{istv} = \beta_1 \cdot Commission_{it} + \beta_2 \cdot Checklist_{st} + \delta X_{istv} + \alpha_i + \gamma_s + \tau_t + \epsilon_{istv}.$$

Y_{istv} is revenue amount (or log revenue amount) for visit v with mechanic i in shop s at time t. $Commission_{it}$ is the commission rate for mechanic i at time t. $Checklist_{st}$ is equal to one if the managerial-control treatment is occurring at shop s at time t, and zero if not. α_i are mechanic fixed effects, which controls for between-mechanic differences in revenue per visit. γ_s are shop fixed effects, which controls for between-shop differences in revenue per visit. τ_t are year-by-week fixed effects to control for any seasonality or other firm-wide time-specific effects. X_{istv} includes car and customer characteristics. Exploiting the high-frequency data at the visit level and including X_{istv} helps to improve statistical precision given the modest number of treatment mechanics. Finally, ϵ_{istv} is an idiosyncratic error term.

Because the managerial-control experiment occurs at the shop level, all of the standard errors account for clustering at the shop-year level. Because these standard-error calculations rely on large-sample asymptotic methods, we show in Section IV that the results are robust to other small-sample robust-inference methods, bootstrapping, and comparing the estimates to those from an empirical distribution of placebo treatments.

B. Effects of managerial control and performance pay on revenue

The main regression results are in Table 5. Columns 1 and 2 show the coefficients on the treatment and commission variables. These baseline models control for time effects with a year-week indicator, shop effects with a shop indicator, and mechanic effects with a mechanic indicator. The unit of observation is the individual visit and the unit of measurement is dollars. As such, the estimated coefficients on treatment and commission represent the difference in the mean revenue per visit within a mechanic during versus before and after the treatment, and before versus after a commission increase.

The results in columns 1 and 2 indicate that the treatment and the commission increases led to more revenue per visit. The coefficient on treatment is 42.45 ($p<0.01$) indicating that the mean revenue per visit at the same shop with the same mechanic is about $42 higher during the treatment period than before or after. The coefficient on the commission rate is 29.38 ($p<0.01$) indicating that the mean revenue per visit at the same shop with the same mechanic is about $29 higher after a one percentage point commission increase. To assuage concerns that these results are not driven be any differences in the characteristics of the cars in the invoices, column 2 includes a comprehensive set of controls about the car itself, including car mileage, age, and make, customer gender, and whether the customer has an account with the shop or not. These controls are unrelated to commission increase or the timing of the treatments. As such, including these controls increases the precision of the estimates but does not change the results (standard errors are about 15 percent smaller).

Because it is helpful to present results in relative terms, and because the log-revenue model fits the data better than the level-of-revenue model, columns 3 and 4 present the same models where the outcomes is the log of revenue rather than the level. As expected, the managerial-control treatment increases revenue per visit by 22.7 percent while a one percentage point increase in the commission rate increases revenue per visit by roughly 13.5 percent. Examining the ratio of these two estimates provides the increase in the commission rate that would be required to have the same effect at the managerial-control treatment. The results across all models show that the treatment increased revenue per visit by approximately the same amount as increasing the commission by 1.7 percentage points ($p<0.01$). Given that the mean commission rate among mechanics receiving a commission is 17.6 percent, the managerial-control intervention had the same effect on revenue per car as increasing the worker commission rate by 10 percent. The 95 percent upper and lower bounds for this ratio are 1.1 and 2.3 percentage points so that adopting the managerial-control intervention had the same effect of revenue per visit as increasing worker commission rate by between 7 and 14 percent.[25]

C. Are managerial control and performance pay substitutes or complements?

Whether managerial control and performance pay are substitutes or complements has implications for the optimal contracting structure. If performance pay is more effective with stronger managerial control and vice versa, firms should adopt an "all-of-the-above" approach. Conversely, if the two approaches are substitutes then firms should only adopt the more cost-effective strategy. From a theoretical perspective either scenario is possible.

Efficiency-wage theories of worker shirking (e.g., Akerlof 1982) yield predictions that contracts with greater monitoring should contain weaker performance pay, and where monitoring

[25] Column (1) of Table 7 shows that operating profit per visit, measured as revenue minus parts and labor costs per visit, gives approximately the same results as revenue per visit.

is costly there should be stronger performance pay. In such models managerial control and performance pay are substitutes because they are competing solutions to the same problem. In our context this could happen if stronger performance pay improves outcomes by inducing workers to conduct more thorough inspections. As such, forcing workers to conduct more thorough inspections via increased managerial control may eliminate the benefit of higher performance pay. In contrast, in a multitasking principal-agent setting, the optimal compensation structure may involve a combination of monitoring and performance pay (Holmstrom and Milgrom 1987, Milgrom and Roberts 1992). Intuitively, one can use monitoring (managerial control) to induce high effort in one task while performance pay can increase effort in another complementary task. In our context, performance pay (commissions) might induce mechanics to search only for high-revenue repairs but not necessarily spend time identifying a more comprehensive set of repairs, including high, medium, and low-revenue repairs. Stronger managerial control, however, may force workers to conduct more thorough inspections and find all repairs, thereby making the marginal hour induced by performance pay more productive.

Because either scenario is possible theoretically, we test whether managerial-control and commission pay are complements or substitutes empirically. In Table 6, we limit the sample to mechanics that are paid primarily on commission and test for whether the interaction between the commission rate and the intervention is positive or negative. In columns (1) we see that the marginal effect of the managerial-control treatment is about 50 percent larger when we limit the sample to these commission mechanics. However, one cannot reject the null hypothesis that the effects are the same for commission and non-commission mechanics. As a further test, we interact the commission rate with the treatment indicator. Column (2) is estimated at the individual visit level and column (3) is estimated at the mechanic-week level. In both models the interaction is positive and statistically significant, indicating that the managerial-control treatment had a larger effect on mechanics with higher commissions. This suggests that an "all-of-the-above" strategy may be optimal in this context.

IV. ROBUSTNESS CHECKS

Before presenting evidence on the underlying mechanisms, in this section we present evidence to establish that the presented estimates reflect real casual effects. We lay out the potential threats to internal validity and address each of them in turn.

A. The commission effect reflects other changes coinciding with the commission increase

The event-study analysis in Section II shows that the timing of the increase in revenue coincides with the timing of the commission increase and is not driven by pre-existing trend differences. However, these results could still be biased if the shop overall had better outcomes coinciding with the commission increases or if the commission increases are a response to improved overall shop outcomes rather than the reverse. Because each shop has several mechanics working per day, we can test for these possibilities directly by using as a control the other mechanics at the same shop at the same time as those that received commission increases. This test is accomplished by including shop-week fixed effects. If shop-specific time shocks correlated with commission increases drive the effect, then including shop-week fixed effects will eliminate the commission effect. Column 5 of Table 5 shows results with shop-week fixed effects and the coefficient on the commission rate is unchanged. Because the managerial-control indicator is absorbed into the shop-week fixed effects, it is not included.

B. The commission effect reflects business stealing from other mechanics at the same shop

Given that commission increases are not random, one might worry that when mechanics receive commission increases they may have greater control over the cars they work on or repairs they conduct, taking away repair work from other mechanics at the shop. If a business-stealing effect is driving the commission effect, then including shop-week fixed effects will *increase* the commission effect. This is because other mechanics at the same shop will have less revenue as some of their usual repairs are transferred to the mechanic with the commission increase. The results in column 5 of Table 5, which include shop-week fixed effects, show no change in the commission estimate.

C. The managerial-control and commission effects reflect shifting across time or mechanics

To ensure that temporal shifting of effort within mechanics does not explain the managerial-control and commission effects (i.e., doing more repairs per car but working on fewer cars overall), we aggregate revenue for all visits for a given worker to the week level and estimate the models on weekly mechanic revenue. The results are in column 6 of Table 5 and are similar to the visit-level effects. One may also wonder whether increases for one mechanic were offset by decreases for other mechanics or decreases in the total number of cars repaired at the shop. We test for this directly by aggregating the revenue across all mechanics to the shop-week level. Because commission level is a mechanic-defined variable, commission level is not included in the regression model. If shifting across workers or servicing fewer cars overall explains the results then estimating the treatment effect on revenue aggregated to the shop would cause the effect to disappear. The results for this test are in column 7. The estimated effect is positive and significant, indicating that a shifting of effort across cars or workers does not primarily explain the effect.[26]

D. The sample is too small for valid statistical inference

[26] While the shop-week effect is modestly smaller than the visit-level effect, the p-value of the difference is 0.28.

Because the number of treated mechanics and treated shops in our sample is small, it is important to show that our estimates are robust to small-sample inference tests.[27] We do this several ways. First, the standard errors throughout are clustered at the shop-year level to account for the possibility that outcomes within a shop are correlated over time. However, as pointed out in Duflo, Bertand, and Mullainathan (2001), clustering the standard errors can lead to an over-rejection of a true null hypothesis. To address this possibility we follow their suggestion and estimate the effects of 1,318 placebo treatments corresponding to shifting the timing of the actual treatment periods by each of an increasing number of days prior to the actual treatment dates, with no overlap with the true treatment dates.[28] Similarly, we estimate the effect of 1,700 placebo commission increases, where for each we chose 17 mechanics at random, assign one percentage point placebo commission increases at a random time for each mechanic, and estimate the model.

The distribution of the placebo treatments and placebo commission increases are in Figure 3. For both variables the actual estimated effect is larger than 97.5 percent of the placebo replications. As a further check on the robustness of our inferences to small samples we follow Camerer, Gelbach, and Miller (2008) who argue that a more conservative test is to use a wild bootstrap clustered t-statistic. For both variables of interest, the estimated t-statistic is larger than 97.5 percent of the wild bootstrap clustered t-statistics.

E. <u>The managerial-control and commission effects are driven by a few mechanics or shops</u>

When the number of treated units is small, one may worry that the results are due to outlier mechanics or shops. We can demonstrate that this is not the case with a permutation test that plots the distribution of the estimated effects when dropping any two mechanics or all

[27] Note that having small numbers of treated subjects is common among labor-market field experiments. For example Shearer (2004) has nine treated subjects and Bloom et. al. (2013) has 11 treated firms (and 14 plants).

[28] We exclude the second treatment period for shop 2, which is only 11 days long, because including it would cause an overlap of the placebo treatments for shop 2 with the actual earlier treatment period treatment for shop 2.

mechanics at any two shops. These results are in Figures 4 (excluding any two mechanics) and A3 (excluding any two shops). All of the permutations yield positive estimated effects, and all are reasonably close to those in the table. Thus, we find that the estimated effects are not sensitive to individual mechanics or shops, and hence are robust to small-sample inference.

F. The interventions reduce customer satisfaction with the firm and consequently future revenue

Given that the majority of customers represent repeat business, one outcome that could undermine the conclusion that the interventions help the firm is that the increased repairs associated with the interventions turn-off customers and reduce repeat business. We test this by estimating the likelihood that any car returns to the same shop within six months in Table 7 column (2). About 47 percent of cars return within six months, and because there is more than nine months after the most recent treatment period in the sample, right censoring is not a significant concern. The linear probability models indicate that neither the managerial-control treatment nor commission increases affect return business.[29]

G. The managerial-control effect reflects customers consenting to more recommended repairs

Because customers only consent to a fraction of the repairs that mechanics recommend, it is possible that the additional revenue under checklist use is due to customers consenting to more recommended repairs, rather than mechanics discovering more repair work. To test this, we hand-entered the recommended repairs for the 1,827 visits during the managerial-control treatments, the one month immediately preceding each treatment, and the corresponding month in the year before each treatment.[30] If the revenue were only due to customers consenting to more recommended repair work, then there will be no increase in recommended repairs. We test

[29] We also estimate a duration model (a Cox proportional hazard model) that uses the full sample accounting for the right censoring that occurs at the end of the sample period, and obtain similar results.
[30] The recommended repairs and charges were in an inaccessible computer format and required hand-entering, which precluded having recommended repairs for the entire five-year period.

for this in Table 7 column (7). The managerial-control effect for the log of revenue for recommended repairs per visit is 0.412 ($p<0.01$). This estimate is larger than that for actual repairs and suggests that mechanics conduct even more thorough inspections than the actual revenue increases would suggest.

V. MECHANISMS

The results thus far indicate that the estimated managerial-control and commission effects can be interpreted causally. For the remainder of the paper, we investigate the underlying mechanisms behind these effects.

A. Are mechanics working on more cars?

One possible mechanism behind the revenue increases is that mechanics are working on more cars. To test this we estimate the mechanic-week-level model (from Section IV) but with the number of visits per week by the mechanic as the dependent variable. The results are in column 3 of Table 7, and indicate that mechanics did not meaningfully change the number of visits per week. The mean number of visits per week is 16.9 and the estimated effects are -0.53 (standard error of 0.95) for the managerial-control treatment and 0.30 (standard error of 0.39) for the commission increases. Thus, mechanics are not servicing more cars.

B. Are mechanics working harder on each car?

Given that mechanics generated more revenue despite not working on more cars, mechanics were either doing more of the same types of repairs per car or shifting to higher revenue repairs per car. We investigate this issue by estimating the effects on the number of repairs conducted per visit (column 4) and on the number of minutes each mechanic worked per week in (columns 5). These two outcomes reveal notable differences. Commission increases did not affect the number of repairs per visit or the amount of time at work, whereas the managerial-control treatment increased the number of repairs per visit by approximately 9 percent ($p<0.01$) and the number of minutes per week by 9 percent or 259 minutes ($p=0.08$). In summary, the managerial-control treatment generated greater worker effort as measured by the number of repairs per visit and the number of minutes worked. In contrast, commission increases had little effect on the number of repairs per visit or minutes per week.

C. Are mechanics doing more expensive repairs?

A remaining possibility is that the revenue increases reflect mechanics shifting to more expensive repairs, especially given the absence of a commission effect on the number of repairs. We test whether the interventions increase revenue by inducing mechanics to substitute from low to high-price repairs. We estimate the effects of the two interventions on the probability of repair for each of the primary repair types, and then conduct a rank-order test, correlating the estimated effects and the mean price for each repair type. If there is a substitution toward more expensive repairs, the correlation will be positive and if there is not, the correlation will be zero.

The results are in Table 8. The table shows repair types ordered by mean repair price for the repair type, and the estimated effects of the managerial-control treatment and the commission increases on the probability of the repair type. Inspection of this table shows that commission increases led to more of the expensive repairs and less of the inexpensive repairs. Indeed the correlation between the ranking of the repair price and the ranking of the commission-treatment effect is 0.85. In stark contrast, the correlation for the managerial-control treatment is -0.04. Thus, the managerial-control experiment led to more repairs overall and the additional repairs were evenly distributed between low and high-price repairs. Commission increases on the other hand led to fewer low-price repairs and more high-price repairs. This suggests that commission pay incentivizes mechanics to focus on high-price repairs at the expense of low-price repairs.

As a final test of this result, we estimate the effect of commission increases on the expected revenue for recommended repairs. If commission mechanics are generating more revenue because they find more expensive repairs, we will see an increase in recommended repairs. Table 7 column (7) show that commission increases had no effect on invoice amount for recommended repairs despite the increase in revenue for actual repairs. This indicates that commission increases caused mechanics to convince customers to agree to more expensive repairs. That is, commission increases seem to lead to a "pushy salesman" effect. This underscores the possible downsides associated with high-powered incentives: they appear only to have caused mechanics to maximize the incentivized output without affecting their behavior related to the primary inputs of effort and thoroughness. While this may or may not have significantly negative effects for the firm, it would have negative implications for social welfare.

D. Is the managerial-control effect working through greater control and monitoring?

The managerial-control treatment appears to induce mechanics to conduct more repairs per car and to spend more time working. The likely mechanism is that the checklist prompts mechanics to check for additional repairs, regardless of the repair price, that otherwise would have gone unchecked. With these more thorough inspections, mechanics find more repairs worth conducting. As a confirmation of whether this "checklist" mechanism is indeed occurring we exploit the fact that two distinct checklists were used by mechanics. The two checklists contain very similar sets of items but the repairs are listing in different orders (Figures 5 and 6 show samples of the checklists). Because mechanics are likely to start at the top of the checklists and work down, possibly being more careful at the top of the list, one might expect that checklists with, e.g., “windshield wipers” at the top will be more likely to induce extra windshield-wiper repairs. We test for this indication of a checklist effect.

We estimate the checklist effect for the items at the top of each checklist, to see if checklist ordering affects repair probability, and hence whether the checklists themselves are indeed dictating inspection behavior. Results are in Table 9. The items in the top section of checklist 1 are windshield wipers (“Wiper blades”) and lights (“Stop Lamps,” “Headlamps,” “Tail Lamps,” “Marker Lamps,” and “License Lamps”). Providing excellent variation, windshield wipers and lights are the last items on checklist 2. Similarly, tires are at the top of checklist 2, but are scattered in various places and not at the top of checklist 1.

If the checklists themselves are dictating inspection behavior, more wiper and lights repairs should occur under checklist 1 than checklist 2. This is exactly what we find. Wipers and lights are significantly more likely under checklist 1 than checklist 2 ($p<0.01$), while tires are significantly more likely under checklist 2 than checklist 1 ($p<0.01$). Furthermore, wipers, lights, and tires are among the least expensive repairs (mean charges of $24, $34, and $44, respectively), further confirming that checklists induce mechanics to address problems that have smaller monetary rewards. In contrast, an increase in commission corresponds to statistically significant *decreases* in wipers, lights, and tire repairs. Note that alignment and suspension components are at the top of checklist 1 (“Springs/Vehicle Height” and “Inner Tie Rod Ends”) and also checklist 2 (“Alignment (2 & 4)” and “Steering/Suspension”), not providing variation, though both show an increase relative to no checklist.

E. <u>Is the managerial-control effect operating through a Hawthorne effect?</u>

As in any field experiment, there is always to concern that outcomes are driven by a Hawthorne effect. That is, the very act of observing workers and collecting data led to improved outcomes. There are many reasons why this is unlikely to drive our results. First, while mechanics were aware that they were being monitored, monitoring was a central component of the intervention. Insofar as being monitored improved outcomes it is likely due to the stakes attached to the monitoring rather than simply being part of an experiment. Second, unlike the well-known Hawthorne experiment, mechanics were never observed by researchers and their day-to-day activities were not interfered with (apart from the checklists). That is, while mechanics were aware that the firm was observing their outcomes, their actions were not being observed. Third, the intervention, the checklist, is a routine practice in much of the industry, and was introduced to mechanics as something the owner would like to try out, rather than being told that they were in a control or treatment group. Lastly, and most importantly we show that the repairs conducted were those that were indicated on the checklists, which suggests a real managerial-control effect and not simply a general improvement in outcomes.

F. Why do mechanics not conduct thorough inspections on their own?

Given the large estimated effects of the managerial-control experiments, one might wonder why mechanics did not conduct thorough inspections (perhaps with a checklist) outside of the treatment periods. In other contexts, researchers have found that certain worker and management practices are not adopted due to lack of knowledge (e.g., Bloom et. al. 2013, Hanna, Mullainathan, and Schwartzstein 2013). However, in our setting moral hazard can explain the results. Specifically, mechanics must exert additional effort to conduct more thorough inspections, yet only receive a fraction of the additional revenue (the commission rate) generated from this effort. For the more thorough inspections to be privately optimal for the mechanics, the extra revenue from the more thorough inspections multiplied by the commission rate must exceed the extra effort cost of the more thorough inspections.

While we cannot directly observe changes in mechanic effort exertion within any given period of time on the job, we can observe changes in effort in terms of time spent on the job. Given that commission mechanics work 47.9 hours per week and receive $830 in weekly pay on average (from Table 2), this converts to $17.33 per hour on the job. Assuming convex effort costs, $17.33 is a lower-bound estimate of the amount required to induce the worker to work an additional hour on the margin, which would make the worker indifferent between incurring the extra time costs associated with thorough inspections and not.

Table 7 shows that mechanics work an extra 259 minutes per week during the managerial-control treatment period and receive an additional $81.49 per week during this period. This converts to a marginal hourly rate of $18.87 for the extra work time during the treatment period. Thus, the additional revenue earned by mechanics is almost exactly the *lower-bound* estimate of what is required to compensate mechanics for the additional time spent on the job using an average hourly pay rate. If conducting more thorough inspections imposes additional costs besides time costs such as greater exertion or conducting more tedious repairs within any given hour, then mechanics may require much more than $17.33 per hour to be indifferent between conducting thorough inspections and not doing so.[31] As such, under realistic assumptions about effort costs, the extra effort by mechanics of conducting thorough inspections is not worth the private costs.

The firm, however, received an extra $497 per mechanic-week during the treatment period, while compensation costs increased by $81 and parts costs increased by $55, leaving approximately $360 in operating profit per mechanic-week from the more thorough inspections (perhaps modestly less due to taxes). Therefore, when inspection thoroughness is observable

[31] Note that the hourly-pay mechanics are paid 1.5 times their hourly rate for any hours worked over 40, suggesting that the mechanics require a significantly higher pay rate to induce them to work the longer hours.

(e.g., via a checklist), the firm could likely compensate workers for the full costs of conducting more thorough inspections and still retain significant additional profits. That is, the more thorough inspections from the monitoring can likely generate a significant Pareto improvement.

VI. DISCUSSION AND CONCLUSIONS

Through a field experiment at an auto-repair chain, we find that increasing managerial control over mechanics significantly increased their productivity. We also estimate the effect of within-mechanic variation in commission rate on productivity and find that the managerial-control intervention increased firm revenue by the same amount as increasing the commission rate by 10 percent. The results indicate that managerial control is a viable alternative to performance pay at mitigating moral hazard. Furthermore, the managerial-control treatment was larger for mechanics that had higher commission rates, suggesting that in this context managerial control and performance pay are complements. The results also support the theoretical prediction that the optimal incentive contract depends on the level and quality of monitoring.

Investigating mechanisms, we find that mechanics under the managerial-control treatment increase revenue through doing more repairs on each car and working more hours each week. In contrast, mechanics that received commission increases increased revenue by substituting away from low-revenue repairs toward high-revenue repairs and getting customers to consent to higher price repairs, with no increase in time on the job or number of repairs conducted. Because this shifting toward more expensive repairs may reflect mechanics exploiting their informational advantage over customers, the result underscore the possibility that pay-for-performance may encourage undesirable worker actions.

We present evidence that mechanics do not conduct thorough inspections because of moral hazard. That is, because mechanics only receive a fraction of firm revenue, the additional compensation for mechanics for conducting a more thorough inspection is insufficient to offset the associated effort and time costs. Our calculations based on the results indicate that a modest transfer of the profits due to checklist use from firm to mechanic could compensate mechanics for their additional costs and achieve a sizable Pareto improvement.

This study demonstrates empirically the ability of increased managerial control to reduce moral hazard, a central theoretical result that had not been tested empirically. The study also provide evidence that increased managerial control can generate complementarities with performance pay, most likely in settings with multiple complementary tasks. Given the widespread emphasis on performance pay as an incentivizing tool, our results suggest that managerial control may be an additional important tool for designing compensation schemes.

REFERENCES

Akerlof, George (1982) “Labor Contracts as Partial Gift Exchange,” Quarterly Journal of Economics, 97(4), 543-569.

Arrow, Kenneth (1985) “The Economics of Agency,” in J. Pratt and R. Zeckhauser (eds), *Principals and Agents: The Structure of Business*, Boston: Harvard Business School Press, 37-51.

Baker, George (1992) “Incentive Contracts and Performance Measurement,” Journal of Political Economy, 598–614.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2005) “Social Preferences and the Response to Incentives,” Quarterly Journal of Economics, 917-962.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2007) “Incentives for Managers and Inequality Among Workers: Evidence from a Firm-Level Experiment,” Quarterly Journal of Economics, 729-773.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) "How Much Should We

Trust Differences-in-Differences Estimates?," Quarterly Journal of Economics, 119(1), 249-275.

Bertrand, Marianne, and Antoinette Schoar (2003) “Managing with Style: The Effects of Managers on Corporate Policy,” Quarterly Journal of Economics, 118(4), 1169–1208.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013) “Does Management Matter? Evidence From India,” Quarterly Journal of Economics, 128(1), 1-51.

Boorman, Daniel (2001) “Today’s Electronic Checklists Reduce Likelihood of Crew Errors and Help Prevent Mishaps,” International Civil Aviation Organization Journal, 1.

Bosk, Charles L., Mary Dixon-Woods, Christine A. Goeschel, and Peter J. Pronovost (2009) “Reality Check for Checklists,” The Lancet, 374(9688), 444-445.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2013) “The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico,” working paper.

Camerer, A. Colin, Jonah Gelbach, and Douglas Miller (2008) “Bootstrap-Based Improvements for Inference with Clustered Errors,” Review of Economics and Statistics, 90(3), 414-427.

Eisenhardt, Kathleen M. (1988) “Agency- and Institutional-Theory Explanations: The Case of Retail Sales Compensation,” The Academy of Management Journal, 31(3), 488-511.

Foster, A., and M. Rosenzweig (1994) “A test for moral hazard in the labor market: contractual arrangements, effort, and health,” Review of Economics and Statistics, 76(2), 213-227.

Freeman, R. B., and Kleiner, M. M. (2005) “The Last American Shoe Manufacturers: Decreasing Productivity and Increasing Profits in the Shift from Piece Rates to Continuous Flow Production,” Industrial Relations: A Journal of Economy and Society, 44(2), 307-330.

Gawande, Atul (2009) *The Checklist Manifesto: How to Get Things Right*, Metropolitan Books.

Glewwe Paul, Nauman Ilias, and Michael Kremer (2010) “Teacher Incentives,” American Economic Journal: Applied Economics, 2(3), 205-227.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein (2013) “Learning Through Noticing: Theory and Experimental Evidence in Farming,” working paper.

Harris, M. and Raviv, A. (1978) “Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement,” American Economic Review, 68, 20-30.

Haynes, Alex B., et. al. (2009) “A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population,” New England Journal of Medicine, 360(5), 491-499.

Hossain, Tanjim and John List (2012) “The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations,” Management Science, 58(12), 2151-2167.

Hubbard, Thomas N. (2000) “The Demand of Monitoring Technologies: The Case for Trucking,” Quarterly Journal of Economics, 115(2), 533-560.

Hubbard, Thomas N. (1998) “An Empirical Examination of Moral Hazard in the Vehicle Inspection Market,” Rand Journal of Economics, 29(2), 406-426.

Jackson, C. Kirabo (forthcoming) “Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence From Teachers.” Review of Economics and Statistics.

Jackson, C. Kirabo, Jonah Rockoff, and Douglas Staiger (forthcoming) “Teachers' Effects and Teacher-Related Policies,” Annual Review of Economics, 6.

Jackson, C. Kirabo, and Henry S. Schneider (2011) “Do Social Connections Reduce Moral Hazard? Evidence From the New York City Taxi Industry,” American Economic Journal: Applied Economics, (3)3, 244-267.

Johnson, Justin P., Henry S. Schneider, and Michael Waldman (2013) “Information Costs, Transactions Costs, and Leasing: Theory and Evidence from the Automobile Industry,” working paper.

Krueger (1991) “Ownership, Agency, and Wages: An Examination of Franchising in the Fast Food Industry,” Quarterly Journal of Economics, 75-101.

Larkin, Ian (forthcoming) “The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales,” Journal of Labor Economics

Lindenauer, Peter K. et. al. (2007) “Public Reporting and Pay for Performance in Hospital Quality Improvement,” New England Journal of Medicine, 356(5), 486-496.

List, John and Imran Rasul (2011) “Field Experiments in Labor Economics,” *Handbook of Labor Economics*, volume 4a, Elsevier B.V., 103-228.

Lucas, Robert (1978) “On the Size Distribution of Business Firms,” Bell Journal of Economics, 508–523.

Mirrlees, J.A. (1999) “The Theory of Moral Hazard and Unobservable Behaviour: Part I,” Review of Economic Studies.

Mishel, Lawrence, and Natalie Sabadish (2012) "How Executive Compensation and Financial-Sector Pay Have Fueled Income Inequality," Economic Policy Institute Issue Brief #331, May 2.

Mullen, Kathleen J., Richard G. Frank, and Meredith B. Rosenthal (2009) "Can You Get What You Pay for? Pay-for-Performance and the Quality of Healthcare Providers," National Bureau of Economic Research, Working Paper 14886.

Pronovost, Peter, Dale Needham, Sean Berenholtz, et. al. (2006) "An Intervention to Decrease Catheter-Related Bloodstream Infections in the ICU," New England Journal of Medicine, 355(26), 2725-2732.

Rosenthal, Meredith B., and Robert. G. Frank (2006) "What Is the Empirical Basis for Paying for Quality in Health Care?," Medical Care Research and Review, 63(2),135–57.

Schneider, Henry S. (2012) "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair," Journal of Industrial Economics, 60(3), 406-433.

Schneider, Henry S. (2010) "Moral Hazard in Leasing Contracts: Evidence from the New York City Taxi Industry," Journal of Law and Economics, 53(4), 783-805.

Shearer, B. (2004) "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," Review of Economic Studies, 71, 514–534.

Thomas Lemieux, W. Bentley MacLeod, and Daniel Parent (2009) "Performance Pay and Wage Inequality," Quarterly Journal of Economics, 124(1), 1-49.

Webb, Natalie J., and James S. Blandin (2006) "Evaluating Executive Performance in the Public Sector," International Public Management Review, 7(1), 98-117.

Wright, P., Kroll, M., Krug, J. A., and Pettus, M. (2007) "Influences of Top Management Team Incentives on Firm Risk Taking," Strategic Management Journal, 28, 81-89.

Yellen, Janet (1984) "Efficiency Models of Unemployment," American Economic Review Papers & Proceedings, 74(2), 200-205.

Table 1: Repairs types and characteristics

	Visits with repair (percent)	Mean revenue ($)	Standard deviation revenue ($)
A/C repair, inspection	2.5	232	271
Alignment, suspension	12.1	285	325
Battery repair, service, inspection	3.8	95	58
Belts, pulley, tensioner	3.2	154	148
Brake repair, fluid, flush, inspection	16.3	311	248
Coolant fluid, flush	3.9	74	65
Engine cleaner, flush	3.0	20	21
Exhaust repair, inspection	9.3	216	219
Filters (air, cabin, fuel, PCV)	4.8	38	33
Fuel cleaner, service	1.4	124	28
Lights	2.9	34	49
Oil change	54.2	25	15
Radiator, hoses, fan, thermostat, water pump	2.4	257	203
Spark plugs, wires, coil, rotor, distributor	2.6	237	159
Tire rotation, repair, balance	9.2	46	112
Transmission fluid, service	1.6	122	66
Windshield wipers	2.2	24	23

Notes: Repair types that occur in at least one percent of visits are included. N=155,049 observations.

Table 2: Summary statistics on mechanics

	All mechanics	Commission mechanics
Commission rate (percent)	10.8	17.6
	[8.7]	[1.9]
Days per week	4.90	4.93
	[1.16]	[1.13]
Visits per week	19.1	18.6
	[10.4]	[8.5]
Revenue per week	3,717	4,034
	[2,412]	[2,268]
Payroll hours per week	47.2	47.9
	[11.9]	[10.8]
Pay per week ($)	798	830
	[360]	[333]
Experience (n visits)	3,706	5,038
	[4,227]	[4,967]
N mechanics	83	51
N visits	147,007	109,099

Notes: "All mechanics" includes mechanics with pay structures of commission, hourly, and flat rate. Twenty-five managers are excluded from the table because they only occasionally conducted repairs, and including them would distort statistics in the "All mechanics" column. Visits for which mechanic information is missing are excluded from the table (approximately five percent of visits). "Experience (n visits)" is the number of visits between 1998 and the current visit that the mechanic conducted at the firm. Standard errors are reported in brackets.

Table 3: Characteristics of treatment and control shops

	Treatment shops	Control shops
N visits per shop-day	11.8	12.0
N mechanics per shop-day	3.57	3.77
Mechanic experience	4125	4578
Mileage per car	94,931	107,687
Fraction visits with oil change	0.568	0.531
Revenue per visit ($)	191	190

Notes: Treatment shops are those in which checklist were used. Control shops are those in which checklists were not used. "Mechanic experience" is the number of previous visits at any of the firm's shops by that mechanic.

Table 4: Description of commission-rate changes

Mechanic identifier	Commission change	Date of change
32	17% to 18%	11/29/2009
71	17% to 18%	8/2/2011
203	16% to 17%	8/14/2011
206	2% to 3%	8/26/2011
302	16% to 17%	8/7/2011
302	17% to 18%	10/21/2012
303	16% to 17%	2/14/2010
402	16% to 17%	11/29/2009
412	17% to 18%	1/18/2009
601	17% to 18%	2/13/2013
602	18% to 19%	7/18/2010
704	16% to 17%	9/4/2011
920	17.5% to 18%	10/8/2008
920	18% to 19%	5/11/2009
920	19% to 20%	3/11/2012
1201	17% to 18%	8/30/2009
1205	0% to 3%	8/19/2012

Notes: The three mechanics with commission rates between 0 and 3 percent are paid primarily on an hourly rate, which is supplemented with the indicated commission rate. The remaining mechanics are paid on commission, with a guaranteed minimum base pay in case of very low commissions that pay period.

Table 5: Estimated models for revenue

	(1) Revenue per visit	(2) Revenue per visit	(3) Log revenue per visit	(4) Log revenue per visit	(5) Log revenue per visit	(6) Log revenue per mechanic-week	(7) Log revenue per shop-week
Treatment	42.45***	40.02***	0.227***	0.202***		0.362***	0.151**
	[15.01]	[13.07]	[0.048]	[0.040]		[0.102]	[0.063]
Commission	29.38***	25.60***	0.135***	0.116***	0.112***	0.191***	
	[6.95]	[6.13]	[0.032]	[0.028]	[0.038]	[0.046]	
Treatment/commission	1.45**	1.56**	1.69***	1.74***		1.89***	
	[0.62]	[0.63]	[0.59]	[0.58]		[0.57]	
Mechanic FEs	X	X	X	X	X	X	
Shop FEs	X	X	X	X		X	X
Year-week FEs	X	X	X	X		X	X
Shop-year-week FEs					X		
Car and customer controls		X		X	X	X	X
Observations	155,177	150,009	152,180	147,351	147,351	9,117	2,843
R-squared	0.108	0.159	0.194	0.262	0.289	0.733	0.659

Notes: "Treatment" is an indicator for visit during the managerial-control treatment. "Commission" is the commission rate of the mechanic in units of percent. "Treatment/commission" is the ratio of the estimates. The models are estimated by OLS. Treatment is omitted from column (5) because the effect is absorbed into the shop-year-week fixed effects. Commission is omitted from column (7) because commission is a mechanic-level variable. The observations in column (7) are weighted by the number of mechanics in that shop-week. Heteroskedasticity-robust standard errors clustered at the shop-year level are reported in brackets. All models include an intercept term. ** and *** indicate statistical significance at the five and one percent levels, respectively.

Table 6: Estimate models of revenue for commission mechanics only

	(1) Log revenue per visit	(2) Log revenue per visit	(3) Log revenue per week
Treatment	0.321***	-0.569	-1.281
	[0.058]	[0.382]	[0.792]
Treatment x commission		0.047**	0.076*
		[0.021]	[0.046]
Commission	0.091***	0.091***	0.147**
	[0.029]	[0.029]	[0.058]
Observations	104,026	104,026	5,846
R-squared	0.205	0.205	0.587

Notes: The sample is limited to mechanics paid primarily on commission. “Treatment” is an indicator for visits during the managerial-control treatment. “Commission” is the commission rate of the mechanic in units of percent. The models are estimated by OLS. Heteroskedasticity-robust standard errors clustered at the shop-year level are reported in brackets. All models include fixed effects for mechanic, shop, year-week, car and customer characteristics, and an intercept term. *, **, and *** indicate statistical significance at the ten, five, and one percent levels, respectively.

Table 7: Estimated models for number of visits, repairs, minutes, and pay

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log of profit per visit	Indicator for return within six months	Number of visits per mechanic-week	Number of repairs per visit	Minutes per mechanic-week	Pay per mechanic-week ($)	Log of recommended revenue per visit
Treatment	0.267***	-0.004	-0.527	0.132***	259.05*	81.49**	0.468***
	[0.050]	[0.014]	[0.950]	[0.0363]	[146.92]	[39.02]	[0.128]
Commission	0.189***	0.018	0.300	-0.010	-2.30	90.60**	-0.020
	[0.042]	[0.012]	[0.391]	[0.0131]	[38.96]	[37.64]	[0.139]
Mean of dep. var.		0.457	16.92	1.440	2940	631.5	
Observations	144,381	42,228	9,434	150,122	5,038	4,956	2,770
R-squared	0.289	0.342	0.726	0.141	0.398	0.737	0.252

Notes: "Treatment" is an indicator for a visit during the managerial-control treatment. "Commission" is the commission rate of the mechanic in units of percent. "Log of recommended revenue per visit" is the estimate quoted to the customer, some of which is sometimes not consented to by the customer. The models are estimated by OLS. Heteroskedasticity-robust standard errors clustered at the shop-year level are reported in brackets. All models include controls for shop, mechanics, customer characteristics, car characteristics, and an intercept term. *, **, and *** indicate statistical significance at the ten, five, and one percent levels, respectively.

Table 8: Rank-order correlation tests of repair-type revenue and intervention effects

	Mean revenue ($)	Treatment effect	Commission effect	Rank of mean revenue	Rank of treatment effect	Rank of commission effect
Brake repair, fluid, flush, inspection	311	0.0359**	0.0204***	1	1	1
Alignment, suspension	285	0.0186	0.0097*	2	3	2
Radiator, hoses, fan, thermostat, water pump	257	0.0036	0.0034*	3	13	8
Spark plugs, wires, coil, rotor, distributor	237	-0.0067	0.0061*	4	16	3
A/C repair, inspection	232	-0.0008	0.0052*	5	15	4
Exhaust repair, inspection	216	0.0125	0.0002	6	8	11
Belts, pulley, tensioner	154	0.0081	0.0022	7	9	9
Fuel cleaner, service	124	0.0144***	0.0041**	8	7	6
Transmission fluid, service	122	0.0061**	0.0020	9	12	10
Battery repair, service, inspection	95	0.0025	0.0049*	10	14	5
Coolant fluid, flush	74	0.0169***	0.0039*	11	4	7
Tire rotation, repair, balance	46	0.0080	-0.0049	12	10	14
Filters (air, cabin, fuel, PCV)	38	0.0073	-0.0037	13	11	12
Lights	34	0.0163*	-0.0073***	14	5	15
Oil change	25	-0.0469***	-0.0485***	15	17	17
Windshield wipers	24	0.0158***	-0.0037**	16	6	13
Engine cleaner, flush	20	0.0231	-0.0135*	17	2	16
Rank correlation between mean revenue and treatment effect:						-0.04
Rank correlation between mean revenue and commission effect:						0.85

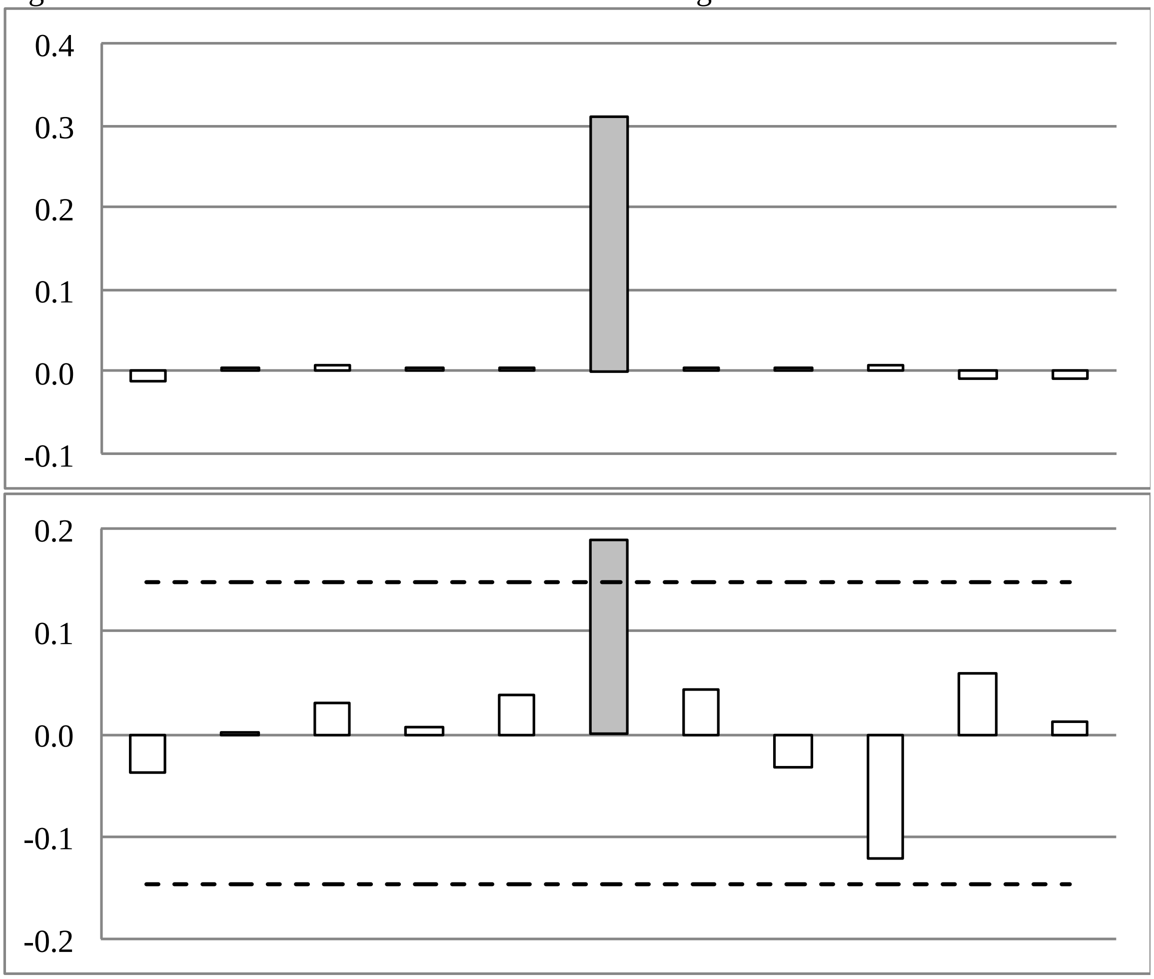
Notes: "Treatment effect" are the estimated effects of the managerial-control treatment on the probability of repair for each repair type. "Commission effect" are the estimated effects of a one percentage point increase in the commission rate. Repairs types comprising that are present in at least one percent of visits are reported. Models are estimated for each repair type individually, and include shop, mechanic, time, customer, and car controls. Results from a rank-order correlation test are reported at the bottom of the table. *, **, and *** indicate statistical significance at the ten, five, and one percent levels.

Table 9: Effect of checklist order on probability of individual repair types

	(1) Lights	(2) Windshield wipers	(3) Tires	(4) Alignment/ suspension
	Top of checklist 1, bottom of checklist 2	Top of checklist 1, bottom of checklist 2	Bottom of checklist 1, top of checklist 2	Top of checklist 1, top of checklist 2
Checklist 1 (treatment)	0.0594***	0.0374***	-0.0237***	0.0185
	[0.0062]	[0.0050]	[0.0050]	[0.0146]
Checklist 2 (treatment)	-0.0026	0.0052*	0.0131	0.0223
	[0.0108]	[0.0030]	[0.0115]	[0.0142]
Commission	-0.0072***	-0.0030*	-0.0061*	0.0116*
	[0.0023]	[0.0017]	[0.0033]	[0.0059]
Checklist 1 - checklist 2	0.0620***	0.0321***	-0.0368***	-0.0038
	[0.0133]	[0.0035]	[0.0125]	[0.0219]
Checklist 1 - commission	0.0666***	0.0403***	-0.0177	0.0069
	[0.0068]	[0.0055]	[0.0125]	[0.0167]
Checklist 2 - commission	0.0046	0.0082	0.0192***	0.0166
	[0.0111]	[0.0068]	[0.0061]	[0.0166]
Observations	155,149	155,149	155,149	155,149

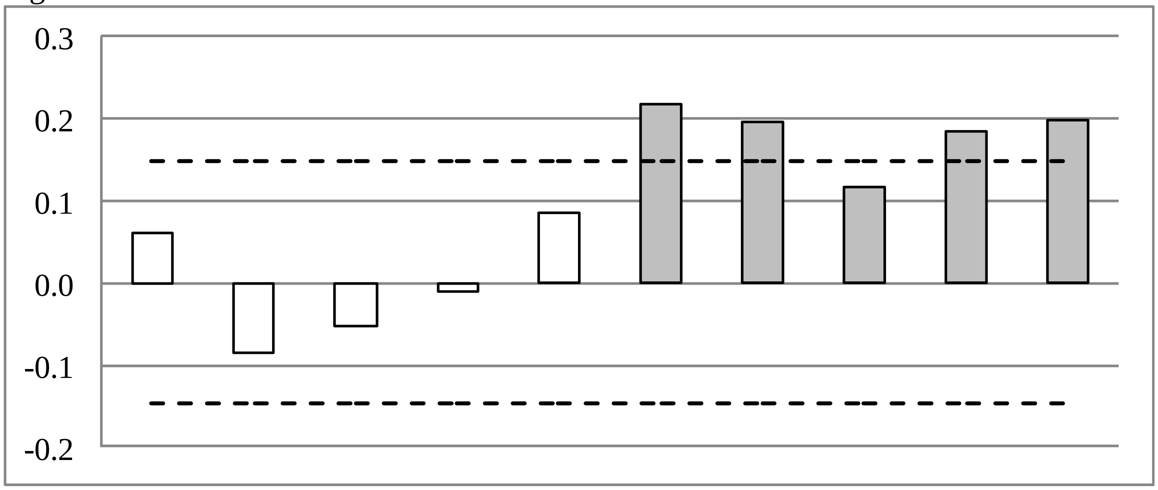
Notes: “Checklist 1 (treatment)” and “Checklist 2 (treatment)” are indicators for visits during the managerial-control treatment with checklist 1 and checklist 2, respectively. “Commission” is the commission rate of the mechanic in units of percent. The models are estimated by OLS and the unit of observation is the visit. The statistics in the bottom panel are the differences in the indicated estimates. Heteroskedasticity-robust standard errors clustered at the shop-year level are reported in brackets. All models include controls for shop, mechanics, time, customer, car, and an intercept term. *, **, and *** indicate statistical significance at the ten, five, and one percent levels, respectively.

Figure 1: Residual checklist use and revenue for managerial-control treatment



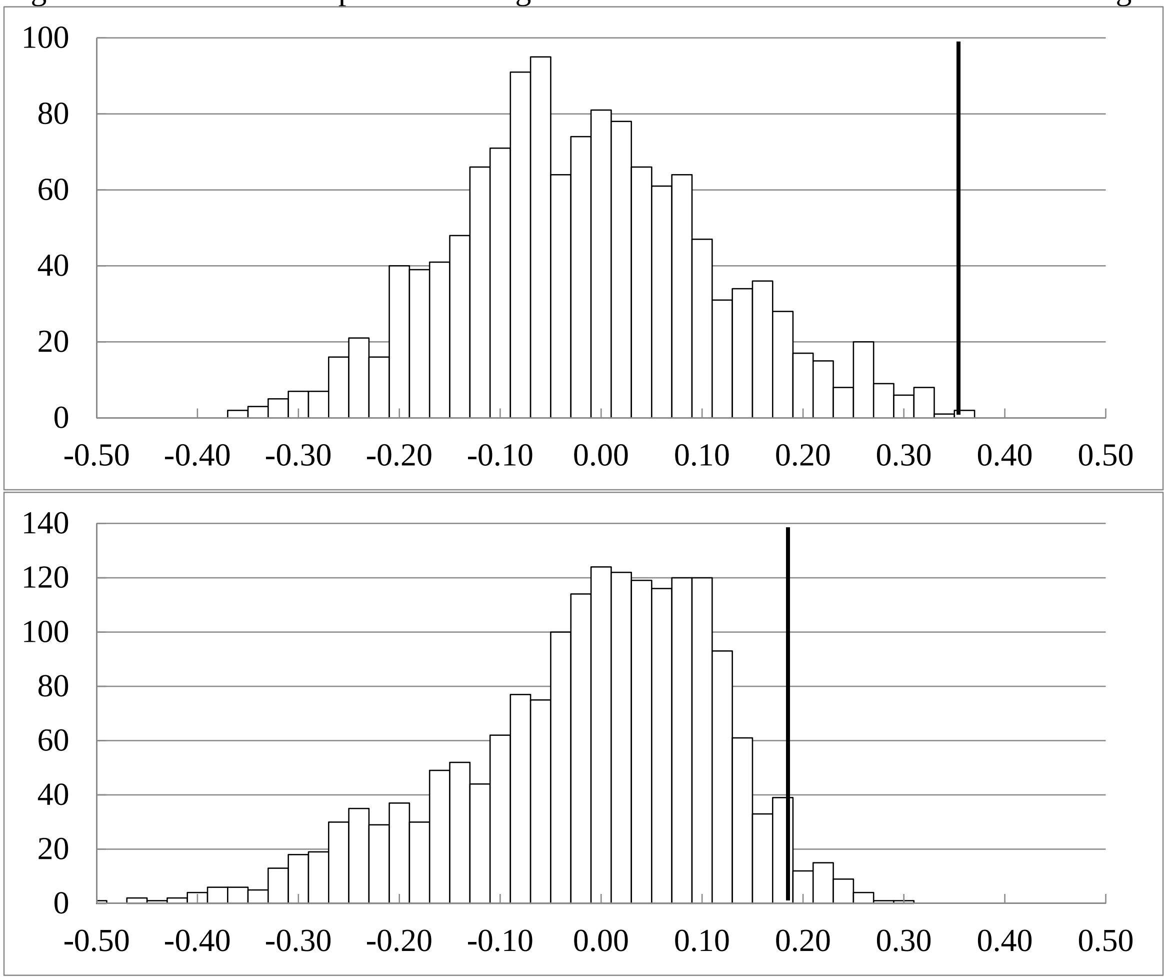
Notes: The bars represent the means of the regression residuals over four-week periods before, during, and after the treatment periods. The dark bar is the treatment period. The unit of observation of the regression is the visit. In the top panel the dependent variable is an indicator for whether a checklist was used in that visit. In the bottom panel it is log revenue for that visit. The bars are normalized so that the mean of the bars for the pre and post-treatment periods is zero.

Figure 2: Residual revenue before and after commission increases



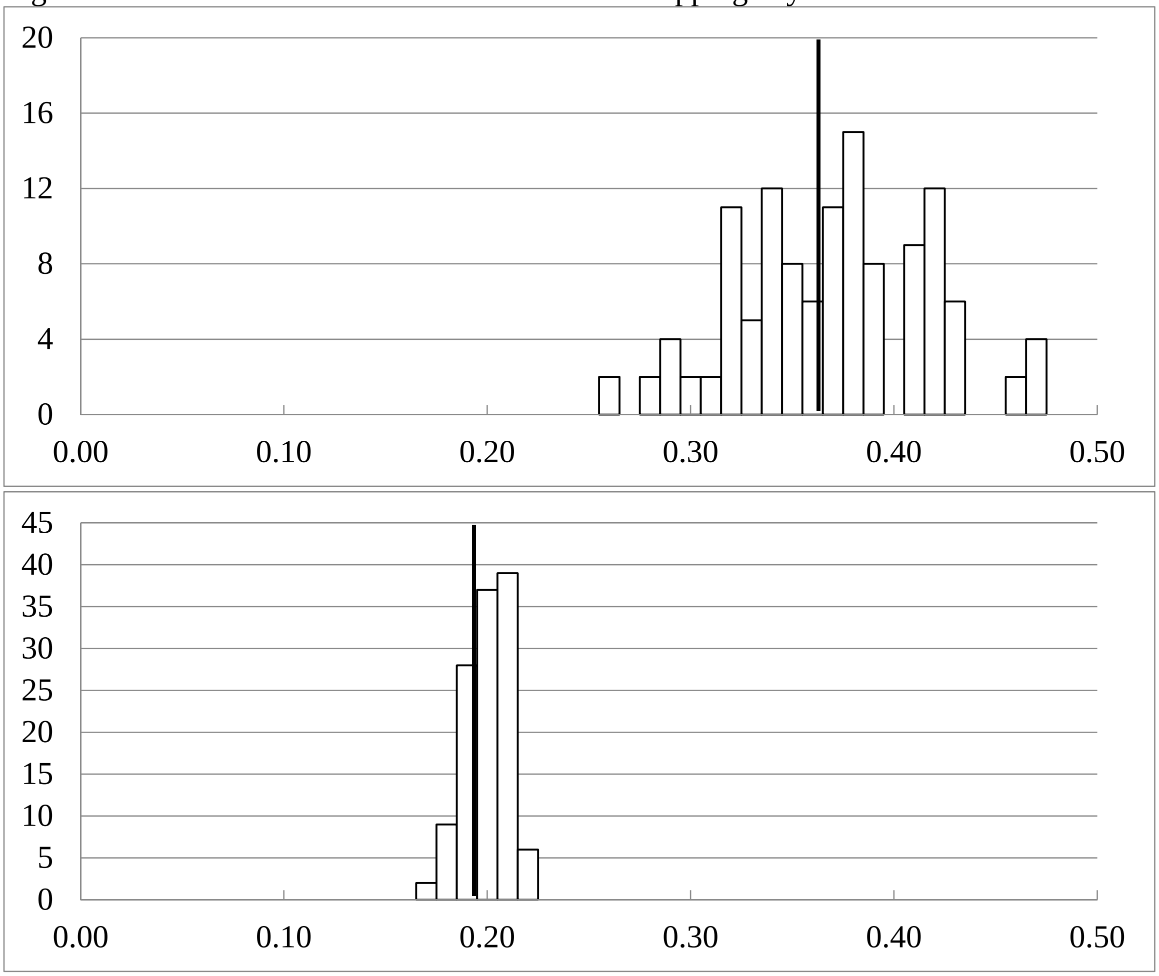
Notes: The bars represent the means of the regression residuals over four-week periods before (light) and after (dark) a commission increase. Log of revenue is for visits for that mechanic. Each dashed line represents two standard deviations from zero, computed from the twelve four-week periods leading up to the commission change. The bars are normalized so that mean of the pre-commission-change bars is zero.

Figure 3: Distributions of placebo managerial-control treatments and commission changes



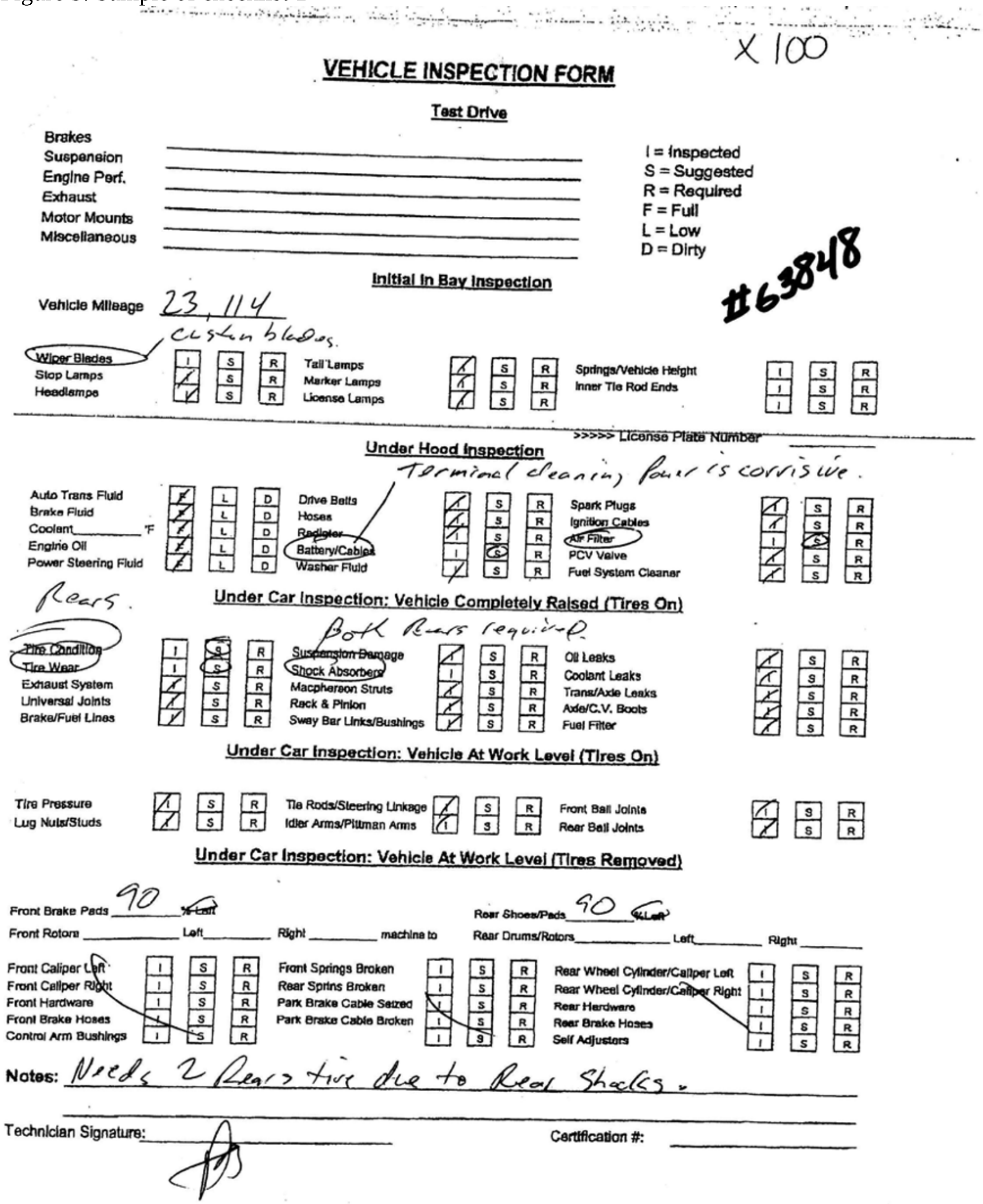
Notes: The top panel shows estimates from the 1,318 placebo managerial-control treatments generated by offsetting the actual treatment periods by an increasing number of days for all days in the sample prior to the actual treatment periods. The bottom panel shows estimates from the 1,700 placebo treatments for 17 commission increases. The vertical lines indicate the estimated effects for the true treatment period (top panel) and true commission rate changes (bottom panel).

Figure 4: Distributions of estimated effects from dropping any two treated mechanics



Notes: The top panel shows estimates of the managerial-treatment effect. The bottom panel shows estimates of the commission effect. All combinations of any two mechanics receiving the managerial-control treatment are dropped and the model estimated. The vertical lines indicate the estimated effects with no dropped mechanics. The models are estimated at the mechanic-week level.

Figure 5: Sample of checklist 1



X 100

VEHICLE INSPECTION FORM

Test Drive

Brakes
Suspension
Engine Perf.
Exhaust
Motor Mounts
Miscellaneous

I = Inspected
S = Suggested
R = Required
F = Full
L = Low
D = Dirty

#63848

Initial In Bay Inspection

Vehicle Mileage 23,114

custom blades.

Wiper Blades	I	S	R	Tail Lamps	I	S	R	Springs/Vehicle Height	I	S	R
Stop Lamps	I	S	R	Marker Lamps	I	S	R	Inner Tie Rod Ends	I	S	R
Headlamps	I	S	R	License Lamps	I	S	R		I	S	R

>>>>> License Plate Number

Under Hood Inspection

Terminal cleaning Power is corrosive.

Auto Trans Fluid	F	L	D	Drive Belts	I	S	R	Spark Plugs	I	S	R
Brake Fluid	F	L	D	Hoses	I	S	R	Ignition Cables	I	S	R
Coolant _____ °F	F	L	D	Radiator	I	S	R	Air Filter	I	S	R
Engine Oil	F	L	D	Battery/Cables	I	S	R	PCV Valve	I	S	R
Power Steering Fluid	F	L	D	Washer Fluid	I	S	R	Fuel System Cleaner	I	S	R

Rears.

Under Car Inspection: Vehicle Completely Raised (Tires On)

Both Rears required.

Tire Condition	I	S	R	Suspension Damage	I	S	R	Oil Leaks	I	S	R
Tire Wear	I	S	R	Shock Absorbers	I	S	R	Coolant Leaks	I	S	R
Exhaust System	I	S	R	Macpherson Struts	I	S	R	Trans/Axle Leaks	I	S	R
Universal Joints	I	S	R	Rack & Pinion	I	S	R	Axle/C.V. Boots	I	S	R
Brake/Fuel Lines	I	S	R	Sway Bar Links/Bushings	I	S	R	Fuel Filter	I	S	R

Under Car Inspection: Vehicle At Work Level (Tires On)

Tire Pressure	I	S	R	Tie Rods/Steering Linkage	I	S	R	Front Ball Joints	I	S	R
Lug Nuts/Studs	I	S	R	Idler Arms/Pittman Arms	I	S	R	Rear Ball Joints	I	S	R

Under Car Inspection: Vehicle At Work Level (Tires Removed)

Front Brake Pads 90 %Left Rear Shoes/Pads 90 %Left

Front Rotors _____ Left _____ Right _____ machine to Rear Drums/Rotors _____ Left _____ Right _____

Front Caliper Left	I	S	R	Front Springs Broken	I	S	R	Rear Wheel Cylinder/Caliper Left	I	S	R
Front Caliper Right	I	S	R	Rear Sprins Broken	I	S	R	Rear Wheel Cylinder/Caliper Right	I	S	R
Front Hardware	I	S	R	Park Brake Cable Seized	I	S	R	Rear Hardware	I	S	R
Front Brake Hoses	I	S	R	Park Brake Cable Broken	I	S	R	Rear Brake Hoses	I	S	R
Control Arm Bushings	I	S	R		I	S	R	Self Adjusters	I	S	R

Notes: Needs 2 Rears tire due to Rear Shocks.

Technician Signature: _____ Certification #: _____

Figure 6: Sample of checklist 2

99 SATURN SL Series 7/20/12
INV# 55882

Inspection Item Required	Condition			Explain/Action
TIRES				
1. Tires - Front	OK	NA	Needs Attention	
2. Tires - Rear	OK	NA	Needs Attention	
3. Alignment (2 & 4)	OK	NA	Needs Attention	
BRAKES AND SUSPENSION				
4 Steering/Suspension	OK	NA	Needs Attention	BOTH OUTER TIE RODS / GREASE BOOTS TORN
5. R&P/Pump/Hoses/Fluid	OK	NA	Needs Attention	
6. Front Brakes	80	% Left	Needs Attention	
7. Rear Brakes	15	% Left	Needs Attention	SHOES / HARDWARE
8. Rotors and Drums	OK	NA	Needs Attention	MACH. NE
9. Hydraulics & Brake Fluid	OK	NA	Needs Attention	
10. Axle Shafts/Boots/U Joints	OK	NA	Needs Attention	
11. Front Struts/Shocks	OK	NA	Needs Attention	
12. Rear Struts/Shocks	OK	NA	Needs Attention	
EXHAUST AND ENGINE				
13. Exhaust System	OK	NA	Needs Attention	
14. Engine/Transmission Leaks	OK	NA	Needs Attention	
15. Differential Fluid Leaks	OK	NA	Needs Attention	
MAINTENANCE AND MISC.				
16. Check Engine Light	OK	NA	Needs Attention	
17. Drive Belts & Accessories	OK	NA	Needs Attention	CRACKS
18. Transmission Fluid	OK	NA	Needs Attention	
19. Cooling System Fluid	OK	NA	Needs Attention	
20. Radiator/Hoses	OK	NA	Needs Attention	
21. Heater Hoses	OK	NA	Needs Attention	
22. Battery & Terminales	OK	NA	Needs Attention	
23. Battery Cables	OK	NA	Needs Attention	
24. Tune-Up Components	OK	NA	Needs Attention	
25. Filters (Cabin/Fuel/Air/PCV)	OK	NA	Needs Attention	
26. Engine Oil (Condition and Level)	OK	NA	Needs Attention	
27. Lights (Brake, tail, turn, etc.)	OK	NA	Needs Attention	
28. Headlights/Adjustment	OK	NA	Needs Attention	
29. Wiper Blades (Front and Rear)	OK	NA	Needs Attention	

30. Factory Scheduled Maintenance Mileage: ________ Due: 3K 7.5K 15K 30K 60K 90K

Miscellaneous Comments:

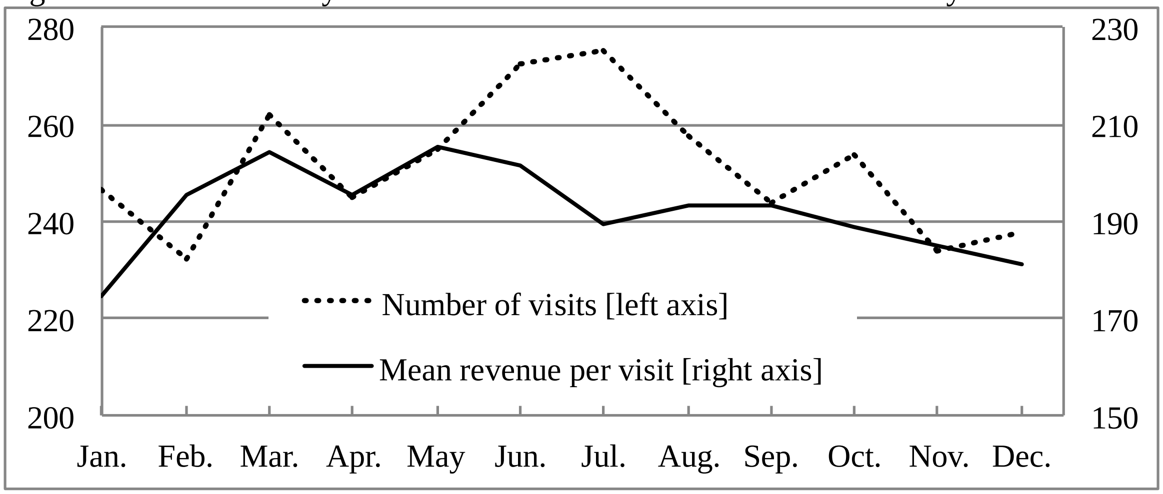
APPENDIX

Table A1: Most frequent 40 line item repairs

Invoice line item	N	Mean charge ($)	CDF
environmental disposal fee	83788	2	0.123
oil filter	83765	8	0.247
top off fluids	70634	4	0.351
shop supplies	57291	18	0.435
5w30 oil	53410	11	0.514
5w20 oil	19886	17	0.543
disc brake rotor	11151	145	0.560
four wheel tire rotation	10255	11	0.575
brake inspection	4899	4	0.582
computerized diagnostic test	4603	59	0.589
air filter	4550	19	0.596
5w30	4113	11	0.602
wiper blade	3892	14	0.608
brake system flush -84032	3067	87	0.612
serpentine belt	2994	78	0.616
5w30 syn blend oil	2897	11	0.621
clean & adjust rear drum brks	2754	29	0.625
exhaust inspection	2723	3	0.629
coolant flush & fill	2688	50	0.633
bg coolant flush kit	2545	30	0.637
ceramic disc brake pads	2480	114	0.640
bleed brakes	2128	25	0.643
exhaust gasket	2075	20	0.646
friction fighter	2049	10	0.649
fuel filter	1997	62	0.652
r134a freon (1/2 lb.)	1941	60	0.655
quickstop brake pad	1897	134	0.658
hardware-gasket	1884	24	0.661
brake shoes	1861	123	0.664
gold extended-life antifreeze	1849	27	0.666
two wheel alignment	1845	45	0.669
reman caliper assy.	1818	164	0.672
machine rotors	1803	41	0.674
mega-tron battery	1736	119	0.677
tie rod end	1690	146	0.679
4 wheel alignment	1666	69	0.682
transmission flush	1664	61	0.684
wagner quickstop	1652	134	0.687
a/c evacuation & recharge	1585	64	0.689
bulbs	1573	13	0.691

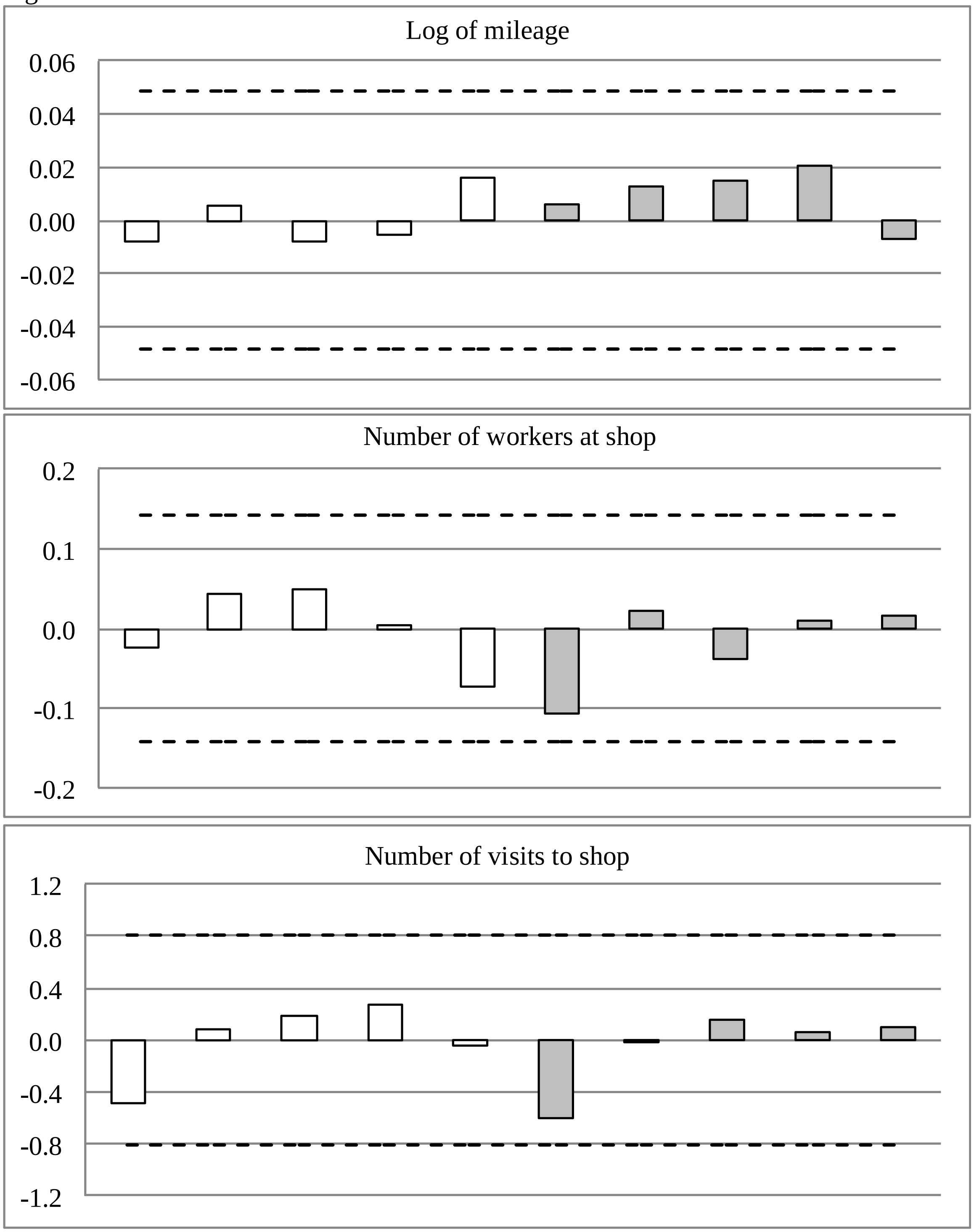
Notes: Invoice line items are the actual itemized repair components listed on the customer invoices and recorded in the data set. Mean charge is the amount charged to the customer for that repair (labor plus parts) before any discounts are applied (discounts are typically applied to the total invoice amount). CDF is the cumulative fraction of line items through that repair.

Figure A1: Mean monthly invoice and number of visits over the calendar year



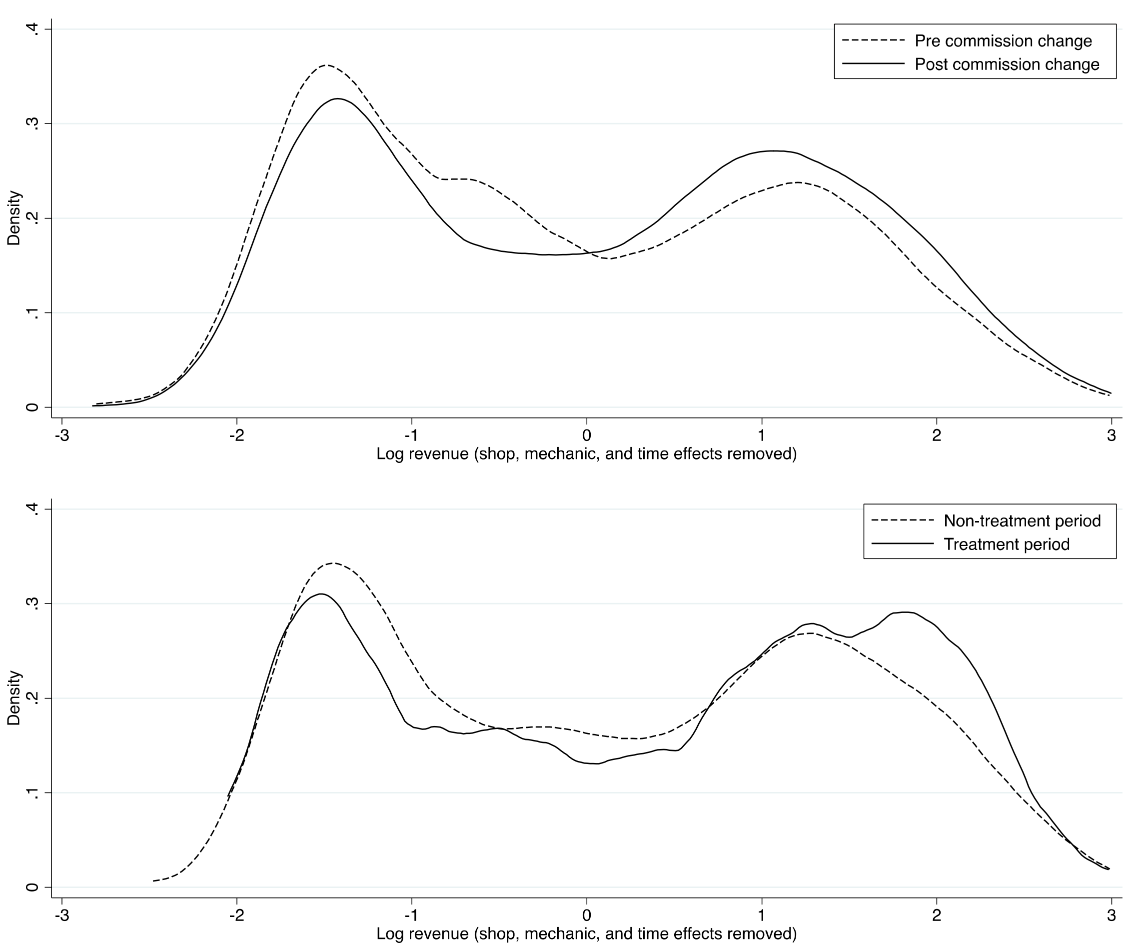
Notes: The data are for the full years 2009, 2010, and 2011. Subsequent data are excluded because treatments occurred at that time.

Figure A2: Residual values before and after commission increases



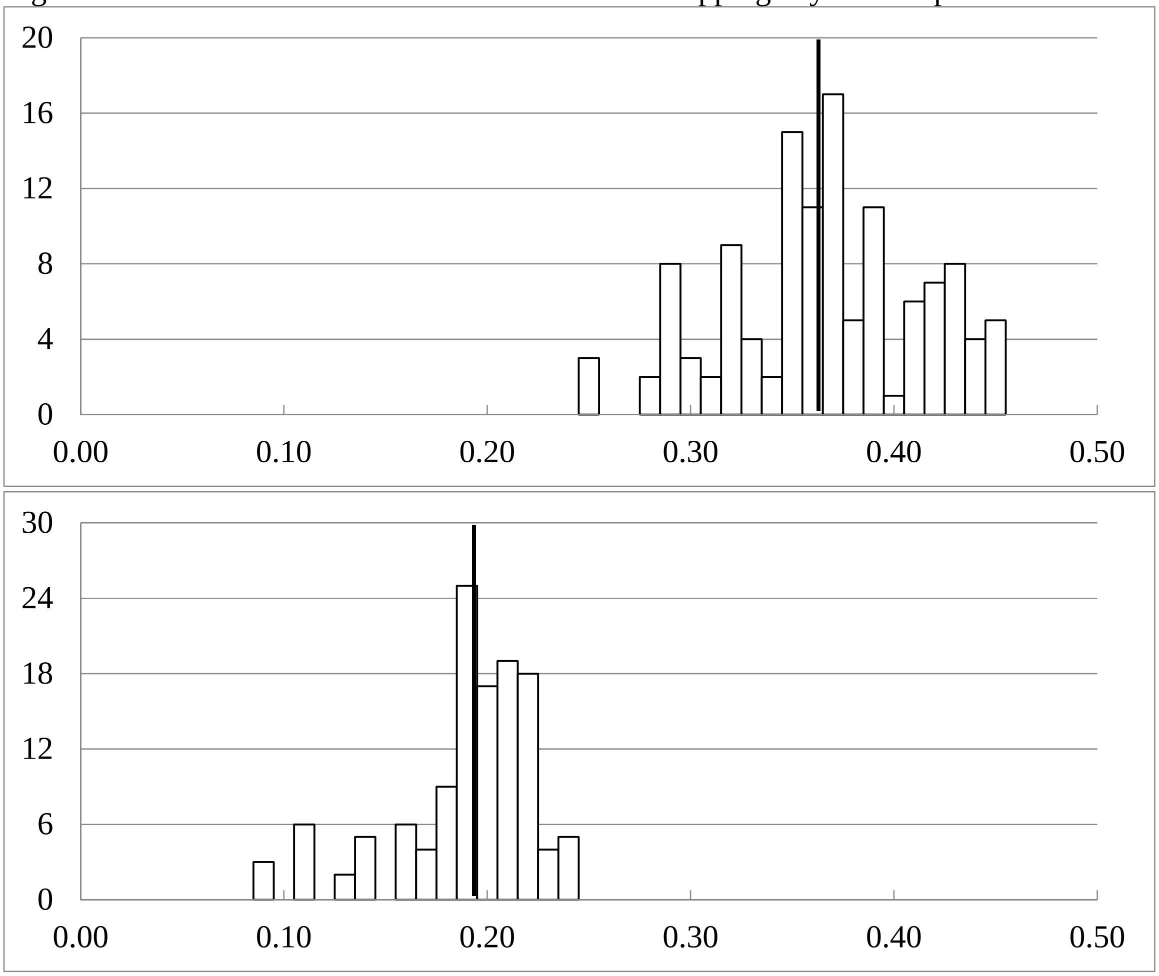
Notes: The bars represent the means of the regression residuals over four-week periods before (light) and after (dark) a commission increase. Log of mileage is for visits for that mechanic. Each dashed line represents two standard deviations from zero, computed from the twelve four-week periods leading up to the commission change. The bars are normalized so that mean of the pre-commission-change bars is zero.

Figure A3: Density plots of log revenue per visit



Notes: The residuals for commission mechanics from a regression of log revenue on fixed effects for shop, mechanic, and year-week are shown. The residuals are plotted with a kernel density smoother with a bandwidth of 0.2. In the top panel, the pre commission change and post commission change curves are for the five four-week periods before and after the commission change, respectively. In the bottom panel, the non-treatment period and treatment period curves are for the five four-week periods before and after the treatment, and the approximately four-week period comprising the treatment periods, respectively.

Figure A4: Distributions of estimated effects from dropping any two shops



Notes: The top panel shows estimates of the managerial-treatment effect. The bottom panel shows estimates of the commission effect. All combinations of any two shops are dropped and the model estimated. The vertical lines indicate the estimated effects with no dropped mechanics. The models are estimated at the mechanic-week level.