The Contingent Effect of Management Practices

Steven Blader
New York University

Claudine Gartenberg
New York University

Andrea Prat
Columbia University

Initial draft: October 30, 2014

This draft: April 14, 2015*

Abstract

This paper investigates how the success of a management practice depends on nature of the long-term relationship between the firm and its employees. A large US transportation company is in the process of fitting its trucks with an electronic on-board recorder (EOBR), which provide drivers with information on their driving performance. In this setting, a natural question is whether the optimal managerial practice consists of: (1) Letting each driver know his or her individual performance only; or (2) Also providing drivers with information about their ranking with respect to other drivers. The company is also in the first phase of a multi-year "lean-management journey". This phase focuses exclusively on changing employee values, mainly toward a greater emphasis on teamwork and empowerment. The main result of our randomized experiment is that (2) leads to better performance than (1) in a particular site if and only if the site has not yet received the values intervention, and worse performance if it has. The result is consistent with the presence of a conflict between competition-based managerial practices and a cooperation-based relational contract. More broadly, it highlights the role of

*We are particularly indebted to Rebecca Henderson who has provided extensive support and advice during this ongoing research program and of course to our partner company for their fantastic collaboration. We would also like to thank Nava Ashraf, Iwan Barankay, Roland Benabou, Nick Bloom, Ray Fisman, Bob Gibbons, Botond Koszegi, Stephan Meier, Raffaella Sadun, Ned Smith, John van Reenen, Christopher Woodruff and participants at the NBER Org Econ Winter Conference, Duke Strategy Conference, SRF Annual Meeting, AEA Annual Meeting and seminars at NYU, Columbia University, and LBS.

intangible relational factors: the optimal set of managerial practices depends on the long-term relationship the company chooses to have with its workers.

1 Introduction

Economists have increasingly focused on management practices as an important explanation for the large observed variation in productivity among firms (Bloom and Van Reenen 2007; Syverson 2011). However, if there are such clear returns to management practices, why do large differences in practice adoption persist across firms, even within narrowly defined industries? One proposed explanation is that poor institutions make adoption costly and also protect inefficient firms from competition (Bloom and van Reenen 2010). While providing a convincing argument for firms in developing countries, this explanation does not address the large observed dispersion in wealthier settings or the substantial within-country variance, where firms presumably operated under similar institutional contexts.

A second explanation, explored by Gibbons and Henderson (2013) and Helper and Henderson (2014), is based on relational contracts. The ability of a firm to introduce a new management practice - and whether that practice is optimal for that firm - is likely to depend on the relational contract that is in place. A relational contract is a non-legally binding understanding between a firm and its employees that typically describes how employees should behave and how the firm will reward the expected behavior (MacLeod and Malcomson 1989; Baker et. al. 2002; Levin 2003). The contract is enforced through repeated interaction and is difficult to change (Chassang 2010; Halac 2012; Baron and Kreps 2013). As a result, similar firms may be governed by very different relational contracts, which can in turn drive differences in adoption of observable management practices.

This proposed effect of relational contracts on responses to managerial practices is difficult to measure for two reasons. First, by definition, relational contracts are difficult to observe. They are implicit informal agreements rather than explicit written contracts exactly because they contain prescriptions that cannot be expressed in a legally binding way. Second, even if they were observable, given that every company has its own relational contract, it would be difficult to disentangle the effect of the relational contract from other firm-specific unobservable factors.

The goal of this paper is to make some progress on this issue by studying a company that is arguably transitioning from one relational contract to another. This complex transition is implemented on a site-by-site basis across the company and is a lengthy process. The company was mid-way through the transition at the time of our experiment and therefore some locations within the firm had experienced the change while others had not. We study how workers in these different locations react to the introduction of the same management practice.

Specifically, we run a field experiment within a transportation company with a large number of sites that all perform a similar function throughout the United States. Our company recently introduced electronic on-board recorder (EOBR) technology that measures the performance of drivers against a route-specific benchmark. The introduction of this technology raises a question about the optimal managerial practice for sharing performance information. In particular, should drivers be made aware only of their individual performance or should performance statistics of all drivers at a given site be posted and made publicly viewable? While both options provide performance feedback to drivers, the latter practice is also likely to spur comparisons and potentially competition between drivers. We worked with the company's management to shed light on this issue by running a randomized controlled field study.

As mentioned above, the company is also in the midst of implementing a far-reaching change to its management culture. They are engaged in a multi-year program to roll out Lean Management. Lean Management ("Lean") is a widespread management philosophy, inspired by the Toyota Production System, centered on teamwork and worker empowerment (Womack et. al. 1990; Holweg 2007). Given the prevailing individualistic culture at our company, a successful implementation of Lean Management requires profound changes in the relational contract with employees across all levels of the organization. Accordingly, the company committed substantial resources to the implementation and a ten-year schedule for the implementation.

This implementation is composed of five phases and, at the time of the experiment, the company was midway through the first phase. Crucially, the first phase involves no change in the work processes of the drivers nor their incentives. It mainly consists of introducing employees to Lean principles and, in particular, how and their emphasis that "continuous improvement" (the organizing idea behind Lean Management) occurs primarily through teamwork, collective effort and the empowerment of front-line workers. This represented a significant shift for employees at our firm, since the previous management philosophy was based on individual supervision and limited delegation. As such, the first phase consisted primarily of setting the stage for Lean Management via formal training, the organization of a "lean team" of front-line employees to model lean culture at each site, and the message to supervisors and managers to begin to change the climate of each site so that drivers lead meetings (rather than managers), raise problems freely and solve them among themselves when possible. For the purpose of this paper, we will refer to this first phase as "the lean intervention" or "lean," even though Lean Management, in its fullest sense, involves many other changes to both formal and informal operating practices (which had not yet been initiated at our company). At the time of the experiment some sites had received the lean intervention while the others had not.

The primary finding in our study is that the effect of posting drivers' performance strongly depends on whether the sites had received the lean intervention. Drivers assigned to untouched sites responded on average positively to the performance postings, improving their fuel efficiency by 4.5% and reducing their idling time and wasted fuel by 1.1% and 1.8%, respectively, relative to the control group. In contrast, drivers at lean sites responded negatively to the individual performance rankings. We record a substantial drop in performance for these drivers, in the form of a 10.7% reduction in fuel efficiency and an increase of 2.5% in idling time and 4.4% in wasted fuel, relative to the control group.

This finding must be interpreted in light of the fact that the lean intervention did not change existing incentives or processes. As such, a researcher who had complete site-by-site 'hard' information about the current managerial practices – but no knowledge of the fact that certain sites had been exposed to the first phase of lean – would have missed a key source of site-level adoption success.

Since the first phase of lean consisted in the announcement of a future change in what the firm will expect from its workforceprimarily consisted of introducing the value of collaboration, teamwork and empowerment—and in communicating the importance these elements will have in the company in the future – our experimental result highlights the importance of 'soft' information on how employees view their relationship with the company. In turn, this intangible relational factor (which we explore in detail later) was by definition under the control of the firm, as the lean intervention was a deliberate choice of the company. The experiment therefore indicates that the optimal set of managerial practices depends on the long-term relationship the company seeks to establish with its workers.

The rest of the paper is devoted to probing the robustness of our result and to understanding the mechanism behind it. Regarding robustness, the biggest challenge in our experimental design is the assignment of the lean intervention across sites. Ideally, we would randomly assign sites to both the performance postings and the lean intervention. However, our study design constrained us to use the pre-existing lean assignment for the experiment: the company's management required a specific four month window for the postings to be rolled out across all the sites, while the initial phase of lean implementation was scheduled to occur over a five year period across all sites, with a minimum of three to six months to complete at any given site. Given this timing mismatch, we instead stratified our randomization of performance postings by whether a site had received the lean intervention at least three months prior to the commencement of the study. We further applied statistical matching between sites with and without lean intervention. Importantly, our results generally strengthened or remained stable after the matching.

Regarding the mechanism, one of the most salient features of this first phase of lean is the Toyota-inspired emphasis on teamwork. Indeed, in a series of interviews we conducted, employees consistently mentioned an increase in "cohesion", "camaraderie", and "respect" following the introduction of lean. Appealing to Benabou and Tirole's (2003) model of intrinsic motivation, we can interpret these changes as a shift in the reference structure that underlies the worker's intrinsic motivation. Rather than maximizing his own overall individual job satisfaction, now each worker also considers his team member's satisfaction. Intuitively, this leads to the prediction that the introduction of lean will reduce the effectiveness of any management practice that relies on competition rather than collaboration between workers. Appendix 1 develops a formal model inspired by Benabou and Tirole (2003) and shows that, with an individualistic reference frame, relative rankings lead to higher performance while, with a collectivistic frame, relative rankings induce worse performance.

To explore this mechanism further, we analyze employee attitudes that the company collected through an annual engagement survey. As one would expect, lean sites score higher on the survey questions that relate to a collectivistic orientation of workers. When we replace the lean intervention dummy with a survey-based collectivist orientation index in our primary triple-differences analysis, we observe a similar pattern of results as our lean indicator. No significant pattern is observed if instead we use a different index of employee attitudes, one that focuses on individual satisfaction with compensation and benefits and thus reflects an individualist orientation. These results indicate that our differential effect are most accurately attributed to collectivistic orientation and not to individual satisfaction.

Two additional tests shed further light on the nature of the backlash we find against relative rankings and provide additional support for our reasoning. First, we find that the triple-difference result extend to second moments. Relative rankings increase performance variance in non-lean sites and decrease it in lean sites. This in line with an explanation based on Benabou and Tirole (2003): with relative rankings, top performers reduce their effort to avoid hurting their teammates' egos, thus leading to a reduction in the variance of performance. Consistent with this, research in social psychology predicts that people in highly collectivist environments have a stronger tendency to develop and comply with shared ways of thinking and acting (Brown and Turner 1981; Tajfel and Turner 1986; Turner 1982) especially when the group experiences a perceived threat (Branscombe et. al. 1999; Spears, Doosje and Ellemers 1997; Tajfel and Turner 1979). These dynamics are especially relevant to organizational contexts (Ashforth and Mael 1992; Blader and Tyler 2009) and thus collectivist environments are likely to show pronounced compression of performance outcomes. Second, we compare the outcomes of two different ranking treatments. In both cases all drivers' scores are publicly posted, but in one case names are withheld (anonymous postings) and in the other case they are revealed (named postings). These two conditions enable us to isolate the effects of competitive dynamics among employees, since they hold constant relative performance feedback and vary only in the identifiability of peers' performance. We find that only the named postings treatment matters. This finding is consistent with social psychological research showing that the competitive, adversarial nature of the rankings should be greatly reduced when one does not know the identities of one's adversaries (Haran and Ritov 2014).

Our study raises a question about the generalizability of findings from randomized control trials within companies, a rapidly expanding method within economics (see Bandiera, Barankay and Rasul 2011 for a survey). Specifically, the results of a field experiment conducted in one organization extend to other organizations with similar observables – or to the same organization at a later time – only if it can be argued that those other organizations have similar relational contracts. To see this, consider our experiment as an example and suppose we accept the conclusion that relative rankings are beneficial to a site if and only if the site is not lean. If we had run the same experiment in 2012 before the beginning of the switch to lean, we would have found overwhelming support for relative rankings. If we had run it after all sites have switched to lean, we would arguably obtain the opposite conclusion. Indeed, although on average we found that relative rankings were weakly beneficial in aggregate (because there are more non-lean sites than lean sites), when the company's management saw our findings they decided not to use relative comparisons because they realized that this management practice went against the relational contract they were attempting to implement. This of course does not mean that experiments within organizations have no external validity. It simply means that intangible factors must be considered when assessing whether experimental findings can translate to another context.

Following a brief literature review, Section 3 provides background information on the research setting and Section 4 describes the nature of our experiment. Section 5 reports the main results, while Section 6 contains the additional tests we perform to further examine the mechanism that underlies our main results. Section 7 concludes.

2 Literature Review

This study relates to a number of literatures within and outside organizational economics.

There are a number of experiments on the effect of relative feedback interventions (see for instance Bandiera et. al. 2012; Barankay 2012; Blanes i Vidal and Nossol 2010; Ashraf et. al. 2014; Bursztyn and Jensen 2015). The results are mixed, with some studies showing an improvement in performance and others showing a decline. While all those works are conducted in setting where organizational culture is held constant, the key contribution of our work is to study the effect of a far-reaching change in the relationship between the firm and its workers.

Most of the experiments in that literature are quite different also because compensation is directly based on the performance measure that forms the object of the relative feedback experiment. Our truck drivers do not have an explicit incentive scheme and promotions are typically seniority based. An exception is Ashraf et. al. (2014), who study health worker trainees in Zambia. All subjects – like ours – receive information about their performance in the training program. Some also learn about their relative performance within the class. Absent additional official recognition mechanism, social comparison worsens trainee performance. This is in line with our findings, but the mechanism is different, because performance worsens *before* trainees receive information about their individual performance, thus indicating a role for anticipatory utility: trainees exert low effort in order to decrease the informativeness of the ranking signal. In our experiment, the change in performance is measured at the point where workers receive information. In line with their explanation, Ashraf et. al. (2014) find that the performance worsening is stronger in the left tail, while, consistent with our relational contract explanation, we find a narrowing of the performance gap. The mechansim we identify in our lean sites also bears a connection to Bursztyn and Jensen (2015). They show that the introduction of a leaderboard in a high school setting led to a 24% decline in student performance. The decline is particularly marked for top performers.

A second conclusion of the paper is that the conjecture by Gibbons and Henderson (2013) and Helper and Henderson (2014) appears to be supported by our evidence: the adoption of management practices requires a complementary relational contract between employees and the firm. Further-more, we expand this logic to find that different relational contracts can shape how workers respond to the introduction of new management practices and that a practice that is beneficial under one relational contract can be detrimental under another one. The link between management practice and intangible characteristics of the firm is present in Ichniowski, Shaw and Prennushi (1997) in their discussion of practice adoption among steel finishing lines. They speculate that differences in practices between lines could be partly explained by differences in levels of trust between labor and management. In support of this explanation, they note that the best practices were observed either in "greenfield" lines or in lines that had been shut down and re-opened under new ownership.

Our work is also related to the theoretical and empirical literature on complementarities between management practices. Milgrom and Roberts (1990, 1995) model complementarities between the elements of complex management systems. Our approach differs in that we analyze the complementarity between a management practice (public/private ranking) and another element that, as we argued above, has no effect on existing processes or incentives. The complementarity we identify is between a management practice and an intangible element related to the way workers perceive their relationship with the company. So, while our findings cannot be understood in a standard complementarity setting, they are consistent with an extension of that approach that includes the expectations of employees regarding their relational contact. Some practices fit with certain relational contracts; other practices fit with other relational contracts. Within management, Levinthal (1997), Rivkin (2000), Porter and Siggelkow (2008) have similarly examined complementarities between firm practices while relaxing strict supermodularity assumptions. Our results in some sense support this approach: we find that lean values and posting of employee performance ranking have negative, rather than positive, interactions with each other.

While there have been a number of empirical tests of the complementarity hypothesis (for example, Ichniowski and Shaw 2003; see Brynjolfsson and Milgrom 2013 for a survey), ours is the first field experiment that explores the complementarity between management practices and non-tangible aspects of the worker-firm relationship. The transportation industry – and in particular the introduction of on-board technology – was studied by Hubbard (2000) and Baker and Hubbard (2003), who documented the presence of crucial design complementarities between monitoring technology, incentive provision, and asset ownership. In our setting, there is no heterogeneity between units in terms of asset ownership and formal incentives.

This paper is also related to research that has noted two important trends in workforce practices over the past three decades. The first trend is the adoption of "innovative" human resource management practices, particularly a trend toward team-based management and group incentives (Lawler et. al. 1995, 2001 and Lawler and Mohrman 2003), perhaps reflecting increased diffusion of Japanese management practices. The second parallel trend has been the increased use of data-driven management, in which firms implement technologies that enable much closer monitoring along (some) key output factors (e.g., Lemieux et al 2009; Cowen 2013). Our paper shows that this two trends, while potentially complementary, have complex interactions that can effect the returns to firms attempting to adopt both.

Finally, the mechanism we identify is related to intrinsic motivation in team problems. A number of empirical studies have examine the effectiveness of group incentives. Contract theory has incorporated intrinsic motivation in incentive problems though a variety of conceptual frameworks (Kandel and Lazear 1992; Koszegi 2013). The closest to the mechanism we identify is Benabou-Tirole's (2003) model of worker type signaling. the switch to lean can be interpreted as a shift in emphasis from an individual reference point to collective reference point. This point is explored more formally in Appendix 1.

3 The Research Setting

3.1 Why the Transportation Industry?

The US transportation industry has several features that make it well suited for research on relational contracts and management practices. Intense competition and well developed information markets (in the form of trade organizations, conferences and consultants) lead firms to rapid adoption of productivity-enhancing technology. Recently, a subset of this technology has provided managers with extensive data and monitoring capabilities, enabling them to implement a broad range of previously infeasible operating practices. In fact, managers are effectively required to do so, given that the technology is only useful insofar as it is effectively integrated into the daily operations of the company. In this sense, we can view these new technologies as a shock to management practices across the industry.

One of these technologies is of particular importance to our research design. Electronic on-board recorder systems (EOBR) record and transmit detailed driving behavior to a centralized database accessible to managers. This database can be used to evaluate, discipline and reward drivers in near-real time. EOBR systems also include terminals installed in truck cabs that display driver performance information and emit audible real-time alarms when driving behavior is out of system bounds.[1]

A second feature of this setting is that the new technology and associated practices can be viewed by drivers as highly intrusive (in fact, at the time of our writing this study, a new technology was announced to install cameras that measure the height of drivers' eyelids to gauge their fatigue[2]). If implemented improperly, firms run the risk of alienating their workforce, which can result in reduced productivity, sabotage and greater union activity. From our discussions with company management, driver acceptance was a primary concern as they decided when and how to roll out new technology. In this sense, the new operating practices can be viewed as complementary to the relational contracts between managers and drivers at these companies.

Related to this point is that the industry has a long history of driver independence: companies have traditionally allowed a high degree of independence to these "last American cowboys," in exchange for long hours and monotonous work.[3] In this sense, the wave of new technology represents a challenge to this tradition and companies are faced with how to handle this transition smoothly.

[1] The industry's interest in EOBRs – also referred to as Electronic Logging Devices (ELDs) – is both the result of regulatory pressure and commercial motives (Koeth 2013). EOBRs are available for many purposes, including safety monitoring, route management, vehicle diagnostics, etc. One of their key potential benefits is to enable "fuel management and fuel use monitoring to improve controls and reduce cost."

[2] http://www.bloomberg.com/news/2014-10-07/droopy-eyelid-detector-one-solution-to-truck-crashes.html

[3] [Trucker culture has been defined by] "the sense of fierce independence, counter-cultural defiance, and unapologetic masculinity...truckers very much valued (and continue to value) not being confined within the four walls of a factory or an office" http://freakonomics.com/2009/02/27/ask-an-economist/)

3.2 The Company

The company at which we conducted this study operates in the less-than-truckload segment of the industry, transporting shipments that are smaller than full truckload freight and larger than individual parcels. At the time of our study, the company employed a substantial number of drivers across sites distributed throughout the US and Canada.[4] Important for our experimental design, most drivers operate local routes and there is little communication between sites. Shipments are picked up and delivered during regular business hours via local routes of less than 300 miles covered by drivers who can serve the same customers over months or years. Intercity shipments are transported between sites via by a minority of "line-haul" drivers, typically on an overnight shift. Because of the difference in shift schedules and the small proportion of line-haul drivers, the threat of cross-contamination between sites during our study is limited. This feature enables us to establish a credible control group and distinct treatment groups in the experiment.

The company was engaged in two major initiatives at the time of our study that we incorporated into our research design. First, beginning in August 2014 and continuing over a four month period, EOBR was rolled out for the first time to all trucks. This rollout represented the first time that company managers had information on individual driver's efficiency and they were sensitive to how the use of this data would be accepted by the workforce. Accordingly, they were open to experimentation on certain practices as a means to decide how to integrate the technology into daily operations.

Second, beginning in 2011 and continuing during our study period, the company was engaged in a decade-long program to change their business culture and operations to conform to Japanese manufacturing practices. At the time of our study, the company had initiated the first of a five phase transition to this "lean" operations model, with a plan to complete the first phase across the remaining sites by the end of 2015. As noted above, the focus of the first phase is primarily on transitioning the culture of the site from an individualistic and hierarchical culture to one of teamwork and empowerment.

[4] The actual number of sites and drivers has been removed for confidentiality purposes, although we discuss the numbers used in our study below.

3.3 The Lean Initiative

The company divided their lean launch initiative into five phases. At the time of our study, only the first phase had been initiated, and only across approximately 35% of the sites. This initial phase was designed primarily to set the stage for the adoption of lean manufacturing processes by instilling in workers an appreciation for the lean manufacturing principles of teamwork and empowerment. Very few formal processes were introduced during this initial phase. Instead, workers at each site went through training on the ideals and principles of lean manufacturing and organized a "lean team" comprised of front-line workers (primarily drivers and others in non-supervisory or management roles). This team underwent additional training and was responsible for instilling "lean values" at the site. This meant, among other activities, having drivers, rather than managers, run meetings and work together to reorganize the community area and dock as they chose. Appendix Figure 2 shows the criteria by which sites are evaluated after completing this initial phase and reflects the emphasis on "soft" changes, such as the nature of the employee-manager relationship and the nature of teamwork at the site. Appendix Figure A3 shows excerpts from interviews with drivers and supervisors on the impact of the lean initiative at their sites. These excerpts indicate that, while workers noticed very few formal changes, they did have a strong sense that the degree of teamwork and management style had both improved as a result.

The timing of this initiative had two advantages for our study. First, the first phase of this transition was primarily focused on changing the prevailing relational contract and involved no changes to formal driver-related workplace practices that could otherwise affect our performance measures. The second through fifth phases did focus on the formal tools side of lean manufacturing, but importantly, none of these phases had been initiated at the time of our study.

Second, 35% of the sites had begun this cultural initiative at least three months prior to the beginning of the study, enabling a meaningful comparison between sites that had undergone the initiative and sites without any culture shift.

Third, after the initial pilot phase, the rollout of the lean initiative was generally decided by the simplest travel schedule of the various managers in charge of training. We consider the rollout, therefore, to be quasi-random for our purposes, in the sense that the rollout schedule is unrelated to the anticipated success of the initiative or other factors that may influence the acceptance of the rank postings.

4 The Experiment

The experiment occurred between August 2013 and July 2014 as EOBR was rolled out throughout the company. We implemented a three by two research design in which we assigned three performance ranking conditions randomly across lean and non-lean sites.

4.1 Performance Rank Postings

We designed two posting treatments in addition to the control group: one in which the driver names were posted next to performance information and one in which the employee IDs were used. In this latter treatment, a driver can identify his own standing and view the distribution within the site, but does not know any other individual's performance, nor do others know his performance. We make use of this latter condition in later sections when we provide evidence in support of the underlying mechanism that we propose drives our main result. Because of the substantial number of sites and the lack of pre-existing outcome data to perform power analyses, we placed equal numbers of sites into each of the three conditions (control, named and IDed postings).

The rankings were posted on a weekly basis, beginning six weeks from the EOBR rollout date for a given site. This timing allowed us to obtain thirty days of pre-measures (we discarded the first two weeks while the systems were calibrated to the trucks). The pre-measures, combined with the control group and lean stratification, enable the triple-differences research design that we describe in Section 5.2.

The postings contain the employee identifier (either driver names or employee IDs, depending on the treatment assignment) and four performance metrics recorded by the EOBR system. These metrics are Gap score, Shift score, Excess idle time and Total fuel lost. We discuss each of these in more depth below in Section 4.4. See Appendix Figure A1 for a sample of the posting. The ranks are determined by Gap Score.

4.2 Sites Included in Field Experiment

Of the total sites at the firm, we discarded more than 20 sites from the sample that launched the lean initiative in 2011, the launch year of the program.[5] In discussions with the company senior management on the need for quasi-random assignment of lean, they mentioned that the earliest sites were selected specifically to be pilot sites, with various reasons for their inclusion. While the company selected approximately ten sites for these initial pilots, we conservatively discarded all sites launched in that first year. We further excluded 36 sites that were scheduled to launch lean during the timeframe of the study, as these locations could not be reasonably be classified as either lean or non-lean. Lastly, we discarded 72 of the sites with postings launched in the first half of the study. Until the mid-project checkpoint in early November, no formal verification process of the rank postings had been instituted and upon further investigation, we learned that there was minimal compliance up to that point. After our discussion, the company instituted a formal process to verify that ranks were posted as required by the experiment guidelines, include weekly photographs of the postings, conference calls, and a shared spreadsheet tracking system.

After these corrections, the experimental sample included more than 5000 unique drivers in 142 sites, 47 in the control group, 50 in the named postings group ("Treatment 1") and 45 in the IDed postings group ("Treatment 2").[6] To construct our driver-day dataset from the EOBR system, we then further removed inter-city routes (defined as routes above 300 miles) and routes with EOBR data that was clearly unreasonable (less than 15 mile routes or MPG$<$1 or $>$15, less than 1% of the sample). This left us with a sample of 330,689 driver-days.

Because the company did not have the managerial bandwidth to reinforce the ranking messaging on an on-going basis, we expected to see some diminishment of any effects of the postings over time. For our main multivariate analyses, therefore, we restricted the windows of the experiment from the thirty days prior to the postings to the thirty days after. We also removed the five days immediately

[5] The actual number of sites discarded in this step has been masked for confidentiality purposes.

[6] The actual number of drivers has been masked for confidentiality purposes.

surrounding the scheduled posting dates, since many of the site managers chose a different day of the work week to post the rankings to coincide with group meetings, rather than on the date specified by the experiment. We were left with a sample of 93,913 driver-days within these narrowed windows that we use in our primary multivariate analyses, although in the appendix we repeat the analyses with the long windows and show that the results are largely replicated, although with somewhat larger standard errors.

Appendix Table A1 contains a summary of the sample construction.

4.3 Outcome Variables

We focus on four outcome variables for this study that capture different aspects of efficient driving performance. Gap score, the primary measure used in the driver rankings, calculates the difference between the average actual and "potential" miles per gallon expended on a given route. The potential miles per gallon is calculated by the EOBR system based on what it considers to be optimal shifting and speed patterns, given weather conditions and route characteristics. Gap score is represented in percentage terms such that, for example, if actual and potental mpgs for a given route are 6.5 and 7.0, respectively, the Gap score would be 7.7 ((7.0-6.5)/6.5*100). A higher Gap score, therefore, represents worse (less efficient) driver performance.

Shift score is the percent of shifting events performed on the route that remains within designated revolutions per minute limits for the engine. For example, if a driver shifts five hundred times on a given route, his Shift score will be 90 if he revs the engine above a designated threshhold during fifty of those shift events. In contrast to Gap score (and the two other measures described below), a higher Shift score represents better driver performance

Excess idle time is a measure of the minutes that an engine idles beyond a designated time period, thereby wasting fuel. This metric particularly captures instances in which the driver allows the engine to idel while making a delivery, counter to company policy.

Lastly, Total fuel lost is an aggregate measure of all the fuel wasted from idling, inefficient shifting, speeding and gearing. As with Gap score, a higher value for Excess idle or Total fuel lost represents worse performance.

Because all four measures are included in the weekly rank postings, we investigate each of them as outcome variables in our analysis, although we focus relatively more attention on Gap score since this is the measure by which the employee ranks are determined.

Table 1 reports summary statistics for the sites in the field experiment. Note that the sample of 142 sites used in the experiment is representative of the full set of 275 sites within the firm, based on observable site characteristics and pre-rank posting driver performance. Within the sample, the Control and Treatment 1 (named postings) groups are statistically indistinguishable, while the Treatment 2 (IDed postings) group appears somewhat different from the other two groups, particularly in the distribution of lean sites and pre-posting driver performance. We control explicitly for these differences in our primary analysis and also perform additional analyses using a matched sample. However, even with these corrections, we interpret results concerning the Treatment 2 group with more caution than those of the Treatment 1 group.

5 Impact of Rankings and Collective Values on Driver Performance

5.1 Preliminary Evidence

Figure 1 depicts the combined response of both lean and non-lean sites to the driver postings, as measured by the each of the four outcome variables. The x-axis is normalized such that week 0 represents the week that rankings were posted at each site, regardless of the calendar date of each posting. The y-axis measures the outcome, with higher values generally signifying worse performance (with the exception of shift score where higher values signify better shifting behavior).

Three results can be drawn from this figure. First, there is no discernable treatment effect: the time histories of the control and the treatment groups appear the same, particularly when comparing the control group to Posting Group 1 (the named postings). Second, performance outcomes generally deteriorate or remain relative constant over time after the EOBR system is launched (at week -6). For example, gap scores for the control group range between 2.1 and 2.2%

for weeks -6 to 0, and increased to 2.2 to 2.3% for weeks 1 to 18. A similar pattern is apparent for Total fuel lost. According to our field interviews, this trend may be attributable in part to the decisions by management not to assign any formal incentives to driving performance, which may in turn have led to reduced focus on the EOBR system over time. Lastly, the final result apparent from Figure 1 is that Posting Group 2 (IDed postings) appears to be different from the two other groups in pre-measures of gap score. Although these differences are not statistically significant, we would have preferred to see closer pre-trends between the two groups. Because of these potential differences, we focus more on the named posting group for our analyses and we draw conclusions from the IDed posting group with more caution and only in conjuction with other tests.

<< Insert Figure 1 about here >>

Figure 2 depicts the performance response to the ranking postings by lean and non-lean sites and show graphically one of the main results of the experiment. For visual clarity, we replace the absolute levels of the performance metrics as shown in Figure 1 with the delta between the treatment and control groups. As such, the 0-line on the Y-axis represents no difference between the two groups, a positive difference (shown by a line above the 0-line) represents worse performance for the treatment group and, conversely, a negative difference (shown by a line below the 0-line) represents better performance (note that, for Shift score the opposite is true).

Three results are apparent from this figure. With the exception of Shift score, these plots shows a clear differences in how lean and non-lean sites responded to the named postings, with the non-lean sites in the named rank group showing relatively better performance than the control group (above the 0-line) and the lean sites showing relatively worse performance (below the 0-line).

Second, the difference between the control and Posting Group 1 (named posting) appears somewhat persistent over the timeframe of the experiment, although near the final weeks of each plot, the performance measures show some convergence. However, two points are important to note regarding this convergence: there is more noise in these final weeks than in the earlier weeks, since only the sites with the earliest ranking rollouts had data that extended this far at the time of the analysis. Also, the company did not reinforce the ranking postings consistently throughout the study period and thus we expected some reversion. As a result, it is not possible to infer whether apparent convergence is an artifact of the data and experiment or a more general finding.

<< Insert Figure 2 about here >>

Overall, this graph shows preliminary evidence for one of the main results of the experiment: that drivers' responses to the performance postings depend on whether their site had transitioned to lean and thus had undergone the shift towards a team orientation. In other words, we interpret these results as showing that drivers' reactions to the postings depends on the relational contract of the site to which he is assigned.

5.2 Intent to Treat Estimates

We now turn to multivariate analyses. We estimate the differential impact of the rankings on lean and non-lean sites using the following triple-differences equation:

$$PERF_{it} = \alpha POSTING\_GROUP_i * LEAN_i * POST_{it} + D'_{it}\beta + e_{it} \qquad (1)$$

where i represents a given driver and t is the calendar date. $PERF$ is one of the four performance outcomes, $POST\_GROUP$ is a vector of two indicator variables, one for each of two possible posting assignments (named or IDed performance postings), and $LEAN$ is an indicator variable that is equal to one if the sites have launched lean at least three months before the beginning of the experiment. $POST$ is equal to one after the assigned date of the posting rollout for the two treatment groups or, equivalently, six weeks after the EOBR rollout for the control group. All pair-wise interactions and individual variables associated with the triple-differences term are also included in the model and α represents the vector of coefficient estimates for all the associated terms. We are primarily interested in the coefficient on the triple interaction itself, which estimates the difference in response to the rank postings of lean and non-lean sites. D_{it} is a vector of control variables that includes the total number of tractors at the site to measure the size of a site, day of week indicators to absorb weekly patterns, lean manager fixed effects, regional fixed effects, and the distance and potential MPG of the route.

We also perform several variations of this analysis to further probe the validity of our initial results. First, we add in driver and date fixed effects to control both for driver traits and seasonality. Second, we create a subsample of the data that matches lean and non-lean sites to account for the quasi-random assignment of lean in the experiment. Lastly, we do an instrumental variables analysis where we instrument actual postings with assigned postings to account for incomplete compliance. We discuss each of these analyses in turn below.

5.2.1 Combined Effect Across All Sites

We begin by estimating the simple intent-to-treat model without differentiating between lean and non-lean sites. Table 2 shows the results of all four performance outcomes, both without and with controls (odd and even columns, respectively). Consistent with Figure 1b, the ranking intervention appears to have no effect. Without accounting for the underlying relational contract at the site, therefore, we might inaccurately conclude that rank postings have no effect on worker performance.

<< Insert Table 2 about here >>

5.2.2 Effect by Lean and Non-Lean Sites

We next estimate the differential impact between lean and non-lean sites. Table 3 shows the results of the intent-to-treat analysis modeled in Equation (1) and is consistent with the plots in Figure 2. In Columns (1) and (2), we see a large, positive difference in the response of lean sites to the named posting treatment. Using the estimates in Column (2), we observe a 13.4% greater average gap score within lean sites with named postings, relative to control, and a 3.8% lower average gap score within non-lean sites with named postings (albeit insignificant).

No similar effect was estimated for the second treatment group that posted the IDed ranks. Consistent with successful randomization, we estimate no statistical difference between treatment groups and control groups (coefficients on Posting Group 1 and 2 and Lean*Posting Group 1 and Lean*Posting Group 2), nor any difference between the control groups in the post-rank posting response (coefficient on Post*Lean). We do, however, see some evidence of overall performance deterioration for the control group (coefficient on Post), possibly reflecting the already-noted observation by company management that the lack of explicit performance incentives may have led to decreased attention paid to the EOBR system over time. Similar results are evident in Columns (3)-(8) for the other three performance outcomes.

Interestingly, the deterioration of the driving performance within lean sites in Treatment Group 1 (named posting) is consistently more significant and larger than the improvement in the non-lean sites. For example, the reduction in Total fuel lost of 5.4% from the Column (8) estimate is larger than the 1.6% deterioration estimated for the non-lean sites.

Also notable here and in subsequent analyses is that the inclusion of control variables does not change the estimates substantially, consistent with successful randomization.

<< Insert Table 3 about here >>

Table 4 repeats the analysis including date and driver fixed effects and the results are similar. Column (1) estimates a 10.7% lower Gap score within Posting Group 1 lean sites relative to control. As in Table 3, we observe no underlying pattern for Posting Group 2, the IDed rank group. The similarity in results between Tables 3 and 4 suggests that the effect of rank postings is due to changes in driver behavior and not due to compositional differences between sites.

<< Insert Table 4 about here >>

5.3 Matched Analysis

One challenge of the study is that lean status is not random. A partial mitigation to this concern is that the company's management indicated to us that, after the 2011 pilot period (not included our sample), the choice to launch lean at a given location was driven by geographic expediency and schedule optimization for the managers involved in the initiative. However, as is evident from statistics in Table 5, this rollout strategy still resulted in some differences between lean and non-lean sites, primarily in the average site size. For our experiment, this size difference would present a challenge in interpreting our main result if larger sites are both lean and likelier to resist rank postings for unrelated reasons. Table 5 also shows other underlying differences between sites, including lower MPG in lean sites, possibly reflecting more urban locations.

For this reason, we construct a matched subsample of sites that adjust for these underlying differences. This subsample includes 78 of the 142 sites in the full experimental sample. The excluded sites include nine of the larger sites in the lean group and fifty five of the smaller sites in the non-lean group. The right side of Table 5 compares the sites across these two groups in this reduced sample and shows that the sites are now statistically indistinguishable across both observable site characteristics and pre-rank posting driver performance.

<< Insert Table 5 about here >>

Appendix Table A2 reproduces the descriptive statistics in Table 2 using the matched sample.

Table 6 reproduces the ITT estimates of Table 3 on the matched sample. The point estimates generally increase and also represent larger percentage standard deviation increases (although they are not statistically different from each other). For example, the Column (2) estimate on Post*Posting Group 1*Lean of 0.1510 represents a 25% standard deviation increase in log gap score, compared to 22% increase based on the estimates in Table 3 Column (2). In general, the results of Table 3 are reproduced and, if anything, strengthened, with the exception of Shift score.

<< Insert Table 6 about here >>

Table 7 replicates the fixed effects analysis of Table 4 on the matched subsample. The results are largely reproduced, again with the exception of shift score.

<< Insert Table 7 about here >>

5.4 Additional Analyses

5.4.1 Potential Changes in Underlying Route Characteristics

We also performed three additional analyses to further rule out potential concerns about our data and experimental design. One such potential concern is whether there are fundamentally different or characteristics of the routes driven in lean and non-lean sites that may explain our observed effects. While we believe that this possibility is remote - it would have to affect only Posting Group 1 (named posting), only at the same six week post-EOBR rollout window as the rank postings and also be orthogonal to the site characteristics on which we based the matched analysis - we perform a placebo test to further rule out this possibility. For this test, we replace our four outcome variables with "potential" MPG. Potential MPG is the system-calculated variable that response to route characteristics and road and weather conditions, but not to driver performance. Therefore, if any route characteristics changed during this period in the lean named-rank group that led to changes in driver performance, we should observe similar patterns in the potential MPG metric.

The results of this analysis are shown in Appendix Tables A3 (on the full sample) and A4 (on the matched sample). Potential MPG shows no changes during this period, while Actual MPG - which is directly related to driver effort - does. It does not appear, therefore, that underlying changes to the routes are driving the results.

5.4.2 Correcting for Compliance

Coordinating the posting rollouts posed a management challenge for the company, particularly since the rankings were rolled out on a weekly basis across 180 sites in 48 states during the busy winter season. As a consequence, approximately 50% of the sites complied with the rank postings during our experiment, even with formal spreadsheet tracking. To account for this selective compliance, we instrumented actual treatment with assigned treatment. The results of this analysis are shown in Appendix Tables A5 and A6 and are stronger than our earlier analyses (including fixed effects and the matched cohorts).

5.4.3 Persistence of Effect

Finally, Appendix Tables A7-A10 repeat the analysis shown in Tables 3-6 without restricting the time windows to the 30 days pre- and post-rank postings. These analyses now include 47 days prior and 207 days after the rank postings. We find that, consistent with some attenuation, the standard errors of the estimates are generally larger, but the effect sizes are close to the narrower-window analysis.

6 How do We Know it is "Collectivistic Orientation" that Matters?

Up to this point, we have simply asserted that the lean intervention created a collectivist-oriented relational contract and that it is this collectivist orientation that drives the different employee response to performance postings. Since relational contracts are, by definition, extremely hard to observe, how do we know that this is the mechanism at work in this experiment?

Although we cannot answer this question definitively, in this section, we present three distinct tests that, taken together, are consistent with our argument. In our first test, we examine the differences in the response between our two treatment groups, the first of which identified the driver performance by name and the second of which identified performance by the anonymous employee ID. In the second test, we look at the effect of rankings on the dispersion of performance within each location. Lastly, we relate driver performance to a proxy measure of "collectivist orientation" based on an employee engagement survey. Each of these tests draws from social psychological research to construct predictions that test the role of collectivist orientation. The results of these tests strongly support our proposed mechanism and its emphasis on the role of collectivist-oriented relational contracts in explaining the effects of lean.

6.1 Named vs IDed Postings

Our reasoning suggests that the effect of the rankings, in both lean and non-lean sites, relies on the identifability of individuals in the rankings. That is, the mechanism we propose for why rankings will have a positive effect in non-lean sites but a negative effect in lean sites relies on the rankings personally identifying each of the individuals. Individual competition to positively distinguish oneself, which we hypothesize motivates drivers in non-lean sites (where an individualistic orientation prevails) (Johnson and Johnson 1989; Stanne et. al. 1999), can only manifest itself if the rankings clearly identify where each individual stands, who is beating whom, who one needs to outdo in order to achieve a higher rank (Anderson and Brown 2010; Gruenfeld and Tiedens 2010; Halevy et. al. 2012; Maholtra 2010) . Similarly, the goal of basking in recognition for good performance and avoiding the shame of being revealed as a poor performer are only relevant in cases where the rankings personally identify each individual (Garcia et. al. 2006). If rankings were instead posted in an anonymous fashion—one in which individuals could only identify their own performance, but not those of their coworkers—then the incentive of positively distinguishing oneself (which motivates those with an individualistic orientation) would dissipate. More generally, a lack of identifiable social comparisons would diminish the relational component of competition, a critical component of competition (Garcia and Tor 2009; Johnson and Johnson 1999; Kilduff et. al. 2010).

Moreover, anonymous rankings should also undermine the mechanism that we propose for the negative effects of rankings among our Lean sites, where a collectivistic orientation prevails. Concerns about erosion of group cohesion and feelings of being connected to others should be less likely to arise when rankings are anonymous, since as noted anonymity strips the rankings of their relational meaning and competitive significance . That is, since specific social comparisons to known peers are not possible when rankings are anonymous, they no longer instigate a competitive, adversarial dynamic. As such, we would expect that the perceived damage to social relations in lean sites would be attenuated since it is unknown who is beating whom, who is disappointing whom, etc.. Therefore, anonymous rankings are less likely to violate the collectivistic-oriented relational contract that is in place in lean sites.

Overall, anonymity should reduce the competitive nature of posted performance rankings, stripping the rankings of the key element that is central to the positive and negative effects of ranking on performance, respectively, in our non-lean and lean sites. If correct, then anonymous rankings should not replicate the pattern of named ranking effects that we present above. If incorrect—for instance, if named rankings have their effect because they convey relative performance feedback (Lazear and Rosen 1981) or even more simply because they convey individual performance feedback—then we would expect that anonymous rankings would demonstrate the same pattern of effects as we found for identifiable rankings.

We utilized this distinction between identifiable and anonymous rankings as a means of exploring the validity of our proposed mechanism. In particular, in our study we included our second treatment group—ID postings—which was likewise included in our random assignment of sites to rankings treatment. In this additional treatment condition, performance was posted in an identical manner to that utilized in our named ranking condition except for one critical difference. Rather than identifying employees by name, we identified them on the ranking charts by their employee IDs.

Consistent with the reasoning above, we find no difference between Posting Group 2 and the control in our triple-difference analysis so far presented.[7]

6.2 Performance Variance

Another test for our proposed mechanism is to examine performance variance, rather than averages. If the difference in driver response to the postings results from a collective relational contract, we predict that named postings should reduce variance in driver performance within lean sites and raise variance within non-lean sites. We predict this pattern for three reasons. First, Benabou and Tirole (2003) predict that top performers will reduce their effort to avoid hurting their teammates egos. Second, as noted above, employees in highly collectivist environments develop more strongly shared ways of thinking and acting, which should compress performance outcomes once they are publicly revealed). Third, consistency in reactions will likely be more common in cases that involve a direct violation to the group's values and norms (Branscombe et al. 1999; Spears et. al. 1997; Tajfel and Turner, 1979), as in the case when named postings are introduced in Lean sites. This is because violations (vs. no violations) are both salient and threatening, thus heightening identification with the group (i.e., merging one's sense of self up with the group) (Ashforth and Mael 1992; Tajfel and Turner 1979) which in turn leads to greater conformity and more normative response patterns (Blader and Tyler 2009; Haslam 2004).

In sum, our speculation that the effects of lean are driven by collective orientation would be supported if performance variance among lean/named rank sites is lower than that of other sites. That is, our proposed mechanism would suggest that there should be greater homogeneity of behavior in lean/named-rank sites. In non-lean sites, individualistic orientations predominate and thus shared understandings about norms and behavior are diminished. Moreover, in lean/no-rank sites, there is no threat present to draw group members inward to the group, its norms, and to each other. Under these circumstances, greater variation among individuals is likely.

In contrast, if performance variance is the same between the lean/named condition and our other conditions, that would call into question our proposed mechanism. In particular, it could either suggest that a) Lean does not breed a collective orientation and/or b) that the rankings are

[7] Tables 3, 5, 6, 7, A3, A4, A5, A6, A7, A8, A9, A10

not interpreted by employees as a violation and threat to the prevailing team-based culture at Lean sites.

Table 8 shows the effect of rank postings on daily performance variance across lean and non-lean sites. Several results are apparent. We find that, overall across both lean and non-lean sites, variance decreases over time. This trend may be due to a learning effect on the part of the drivers or improved instrument calibration. Second, we note that this decreased variance does not occur in non-lean sites with named ranking postings; that is, relative to the control groups, variance increases in non-lean sites with named postings. In contrast, in the lean sites, variance reduces between the control and Posting Group 1 (named postings).

Lastly, we also observe no effect of IDed rank postings on variance in either the lean or the non-lean sites. In terms of economic magnitudes, the difference in response between lean and non-lean sites ranges from 25% of a standard deviation of Log(Gap score) variance to 45% of a standard deviation for Log(Excess idle time). In sum, these results are consistent with our reasoning that collective orientation will compress performance once performance is revealed, while individualistic orientations will increase dispersion.

<< Insert Table 8 about here >>

6.3 Engagement Survey Responses

Our third test of the proposed mechanism takes the most direct approach, in which we use responses to the company's annual engagement survey. This survey was conducted across 45 sites in July 2014 with 564 driver responses. From the survey responses, we created a direct measure of collectivistic orientation as well as a comparison measure of individualistic orientation (specifically, instrumental assessments/judgments of the individual outcomes that one receives from the organization). We compare the effects of these two measures to test our prediction that lean specifically affects drivers' collectivistic orientation, but not other orientations or judgments about the organization, and that this in turn drivers' responses to rankings.

If collectivist orientation drives, we should find the following results using the survey data: a) Lean sites should score significantly higher on employee responses on our collectivist orientation index, and b) the collectivist orientation index should produce a similar pattern of results as our lean indicator in our primary triple-differences analysis. Together, these findings would substantiate two of the most critical components of our proposed mechanism, since they would support our arguments that lean does indeed facilitate collective orientation and that collective orientation underlies the Posting * Lean interaction that we find in our focal analysis. In other words, the survey data enables us to more directly and precisely examine (in a subset of our data) whether the factor that is most central to our proposed mechanism relates, in accordance with our predictions, to both its antecedent (Lean) and its consequences for performance. We also conducted the same analyses using the individualist orientation index. Doing so provides a counterpoint for interpreting our findings, enabling us to determine if a) instrumental individual concerns, rather than collectivistic orientation, better explain our effects and b) if a factor common to both indices—e.g., overall satisfaction—is the factor that actually underlies our findings.

Table 9 shows the relationships between lean sites and our collective and instrumental engagement measures. This table shows that lean sites are associated with higher engagement scores on both our collective and individualist orientation measures.

<< Insert Table 9 about here >>

Since this is a cross-sectional analysis, we may be concerned that underlying differences between lean and non-lean sites drive the differences engagement survey responses. Accordingly, Table 10 shows the same analysis as the previous table using a sub-sample of survey responses in which lean and non-lean sites were matched by size, region, and driver race, age and tenure. The results attenuate somewhat but are statistically the same as the unmatched sample. In particular, the association between lean and collective orientation remains robust.

<< Insert Table 10 about here >>

Table 11 shows the key result of this analysis: a higher score on the collectivist index is associated with a more negative response to named postings, while no such response is observed using the instrumental/individualist index. Figure 3 divides the collective and instrumental indices into deciles and plots each cohort's response to the named postings, relative to the lowest decile. We can see from this figure an increasingly negative response to named posting as collective decile increase, while no such pattern is observed for across the instrumental deciles.

<< Insert Table 11 about here >>

<< Insert Figure 3 about here >>

One cautionary note about this analysis is that there appears to be some pre-treatment differences between drivers with high collective orientation in Treatment Group 1 and the control group (see the Treatment Group 1*[Category] coefficient in Table 11), so we cannot rule out absolutely that our observed patterns are not driven by underlying differences between these two groups. Mitigating this concern is that Table A15 shows that the magnitude of the treatment response appears unrelated to the degree of pre-treatment differences between these two groups.

From this last test, our results indicate: a) a pattern for the collective orientation index that is highly consistent with and supportive of our proposed mechanism, b) a pattern for the instrumental index, an index tied more closely to individualistic orientation, that is not consistent with and thus not likely to explain our primary findings, and thus c) suggestive evidence that it is not something common to both indices—e.g., a generalized sense of satisfaction—that drives our results but rather something unique to collective orientation.

The three preceding tests are supportive of the explanation that lean affects collectivist orientation, which in turn influence the drivers' responses to the rank postings. Notably, the tests are entirely independent of one another yet converge in their support for our emphasis on collective orientation. Any alternate proposed mechanism would need to explain these three observed patterns, each derived from a very different analysis, and thus these tests provide strong support for our reasoning.

7 Conclusion

In this study, we randomized the posting of employee performance rankings across a company that was midway through a costly, multi-year process of altering its relational contract with its employees. Employees working in locations with the original contract responded positively to the performance postings, with their performance improving 2-4% relative to the control group (depending on the performance measure). In contrast, employees in the sites with the newer contract responded negatively, with their performance declining 4-13% relative to the control group.

It appears that these different responses are driven by the individualistic orientation of the initial contract and the collectivist orientation of the new contract. This new contract is based on the Toyota Production System which emphasizes the value of teamwork and cooperation, as well as the subservient role of management whose primary task is to enable the front line workers. This result can be understood within Benabou and Tirole's (2003) model of incentives and prosocial behavior (a formal model is available in Appendix I). Research in social psychology has found that employees respond poorly to perceived inconsistencies in leaders' messages. Our findings support this result, with the posting of individualistic performance rankings representing a violation of the collectivist contract rolled out by the company.

The main contribution of the present paper is to show that the success or failure of a management practice depends on underlying conditions at the firm. These conditions include not just the environment in which the firm operates in and the presence of other management practices, but also on the type of long-term relationship that the firm chooses to establish with its employees. A company who is considering adopting a new practice should ask itself not just whether this practice worked in similar firms, but more specifically whether this practice worked in firms that have a similar relational contract with their employees. This result highlights the importance of measuring not just management practices but also how workers perceive the relationship with their employer.

We have several directions for future research. The timing of this study did not permit us to simultaneously randomize the implementation of both the management practice and the relational contract, since the practice had to be rolled out over the period of three months at the end of 2013 and the initiative to alter the firm's relational contract began in 2011 and required a minimum of five years to be implemented fully. To accommodate this timing mismatch, we stratified the randomized ranking postings by the lean status of each location. In later studies, we aim to randomize the rollout of the relational contract itself, a much lengthier and complex process that should allow us to make more definitive statements about the direct impact of relational contracts on employee productivity. Aside from this direct effect, a further area to explore is the process of altering these relational contracts themselves. Specifically, we would like to understand the factors that determine differences in adoption success. Finally, beyond single firm studies, we would like to extend this research across firms, industries and geographies.

8 Appendix 1: A Model of Complementarities between Lean and Relative Rankings

This stylized model analyzes the effect on employee behavior a change in management practices (the adoption of relative performance rankings) and a change in the company's management philosophy (the adoption of lean), which in turn changes the reference points of workers. The goal of the model is to make predictions on how employee behavior changes when the two practices are introduced either separately or jointly.

In the case of relative performance rankings, the company makes individual performance in a certain activity observable or salient to all workers. The main idea here is that – absent other considerations – performance comparisons make high-performers happier and low-performers less happy. To capture that, let u_i be the direct job satisfaction of worker i. We assume that

$$u_i = y_i + b\bar{y}_{-i} + \rho\left(y_i - \bar{y}_{-i}\right) - \frac{1}{2}cy_i^2;$$

where: y_i is the performance of agent i, $\bar{y}_{-i}$ is the average performance of the rest of the n-person team, namely

$$\bar{y}_{-i} = \frac{\sum_{j \neq i} y_j}{n-1};$$

ρ is a parameter that captures the observability/salience of performance rankings; and c is a cost parameter. Thus, direct job satisfaction consists of four terms:

- Absolute individual performance (y_i).
- The absolute performance of the teammates, though a direct effect ($b\bar{y}_{-i}$), because the agent may care directly about the team output.

- Relative performance ($\rho(y_i - \bar{y}_{-i})$), whose strength depends on how observable/salient rankings are. This can be rationalized as a reduced form of Benabou and Tirole's (2003) model of worker type signaling
- Cost of effort ($\frac{1}{2}cy_i^2$), reflecting the assumption that high performance requires more work.

For lean management, we recognize that it constitutes a complex set of practices that affect worker behavior through multiple channels. However, in the experiment under consideration employees are involved with the first stage of the "lean journey", Bronze. This stage is mainly concerned with mastering the principles of lean management. In particular, employees become familiar with the "Cultural Enablers," which consist of two principles: respect and humility. As Toyoda (1950) put it: "Humility is considered the quality of being modest, unassuming in attitude and behavior. It can also be taken as a feeling or showing respect and deference toward other people." The spirit of humility and respect aims to induce employees to shift from a focus on individual outcomes to collective outcomes.

We capture the effect of the adoption of lean management in a very basic model of interdependent preference model (see Sobel 2005 for a survey). The shift from individual job satisfaction to team job satisfaction is represented as an increase in the importance of the reference group, which in this case is the team the worker belongs to. Namely, recalling that u_i is the direct job satisfaction of agent i, we define U_i as the overall job satisfaction of i and we assume that it depends on his own direct satisfaction but also on that of his coworkers:

$$U_i = (1 - \lambda) u_i + \lambda \bar{u}_{-i},$$

where: U_i is overall job satisfaction; λ is a parameter that captures the extent to which the principle of humility has been internalized by employees (with $\lambda = 0$ being pure individualism and $\lambda = 1$ representing absolute selflessness); and $\bar{u}_{-i}$ represents the average direct utility of the other agents, namely

$$\bar{u}_{-i} = \frac{\sum_{j \neq i} u_j}{n - 1}.$$

The structure of the present model parallels that of the model used in Bandiera, Barankay, and Rasul (2005), where each worker puts some weight on his own payoff and some weight on the payoffs of his or her coworkers.

Now that we have a model that encompasses the introduction of performance rankings and/or lean management, we are ready to characterize the effect of the two practices on employee performance:

Proposition 1 *In equilibrium:*

(i) Without lean management ($\lambda = 0$), the introduction of rankings has a positive effect on agent performance;

(ii) Without rankings ($\rho = 0$), the introduction of lean management has a positive effect on agent performance;

(iii) There is a negative complementarity between lean management and rankings:

$$\frac{\partial^2 \hat{y}_i}{\partial \lambda \partial \rho} < 1$$

(iv) If the presence of lean management is sufficiently strong (λ is large), introducing rankings worsens agent performance.

Proof. The overall job satisfaction of agent i is given by

$$\begin{aligned}
U_i &= (1-\lambda)\, u_i + \lambda \frac{\sum_{j\neq i} u_j}{n-1} \\
&= (1-\lambda) \left(y_i + b\bar{y}_{-i} + \rho\,(y_i - \bar{y}_{-i}) - \frac{1}{2} c y_i^2 \right) + \lambda \frac{\sum_{j\neq i} \left(y_j + b\bar{y}_{-j} + \rho\,(y_j - \bar{y}_{-j}) - \frac{1}{2} c y_j^2 \right)}{n-1} \\
&= (1-\lambda) \left(y_i + b \frac{\sum_{j\neq i} y_j}{n-1} + \rho \left(y_i - \frac{\sum_{j\neq i} y_j}{n-1} \right) - \frac{1}{2} c y_i^2 \right) \\
&\quad + \lambda \frac{\sum_{j\neq i} \left(y_j + b \frac{\sum_{k\neq j} y_k}{n-1} + \rho \left(y_j - \frac{\sum_{k\neq j} y_k}{n-1} \right) - \frac{1}{2} c y_j^2 \right)}{n-1}
\end{aligned}$$

Hence, the marginal effect of a performance increase on agent i's overall satisfaction is given by

$$\begin{aligned}
\frac{dU_i}{dy_i} &= (1-\lambda)(1+\rho-cy_i)+\lambda\frac{\sum_{j\neq i}\left(b\frac{1}{n-1}+\rho\left(y_j-\frac{1}{n-1}\right)\right)}{n-1} \\
&= (1-\lambda)(1+\rho-cy_i)+\lambda\frac{\sum_{j\neq i}\left(b\frac{1}{n-1}-\rho\frac{1}{n-1}\right)}{n-1} \\
&= (1-\lambda)(1+\rho-cy_i)+\lambda\frac{b-\rho}{n-1}
\end{aligned}$$

Yielding first-order condition

$$\hat{y}_i = \frac{1}{c}\left((1+\rho)+\frac{\lambda}{1-\lambda}\frac{b-\rho}{(n-1)}\right)$$

Hence

$$\begin{aligned}
\frac{\partial \hat{y}_i}{\partial \lambda} &= \frac{1}{c}\frac{1}{(1-\lambda)^2}\frac{b-\rho}{n-1} \\
\frac{\partial \hat{y}_i}{\partial \rho} &= \frac{1}{c}\left(\frac{(1-\lambda)(n-1)-\lambda}{(1-\lambda)(n-1)}\right)
\end{aligned}$$

and therefore

$$\begin{aligned}
\left.\frac{\partial \hat{y}_i}{\partial \lambda}\right|_{\rho=0} &= \frac{1}{c}\frac{1}{(1-\lambda)^2}\frac{b}{n-1} > 0 \\
\left.\frac{\partial \hat{y}_i}{\partial \rho}\right|_{\lambda=0} &= \frac{1}{c} > 0 \\
\frac{\partial^2 \hat{y}_i}{\partial \lambda \partial \rho} &= -\frac{1}{c}\frac{1}{(1-\lambda)^2}\frac{1}{n-1} < 0
\end{aligned}$$

and

$$\frac{\partial \hat{y}_i}{\partial \lambda} < 0 \text{ if } \rho > b$$

■

The main result of the model is point (iii). There exists a negative complementarity between the two management practices under consideration: lean management and employee rankings. The effect of rankings on performance becomes lower as lean management becomes more pervasive. This is because lean management makes workers more aware of the “ego bashing” effect that a high individual performance has on other team members. As they care more about team spirit, they are less inclined to ruin it by over-performing and making everybody else feel bad.

The other results are easy to understand once (iii) is in place. Absent lean management, rankings improve performance because agents care about their relative performance with respect to their colleagues (point i). Absent rankings, lean management enhances performance because it makes agents value more their contribution to the team and hence it makes them more willing to work harder (Point ii). Finally, the negative complementarity between lean and rankings means that if the lean intervention is sufficiently strong the effect of the introduction of rankings on performance must be negative (point iv).

References

[1] Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. "Awards Unbundled: Evidence from a Natural Field Experiment." Journal of Economic Behavior & Organization 100 (April 2014): 44–63.

[2] Baker, G, R Gibbons and K Murphy, "Relational Contracts and the Theory of the Firm," The Quarterly Journal of Economics Vol. 117, No. 1 (Feb., 2002), pp. 39-84.

[3] Bandiera, Oriana, Iwan Barankay, and Imran Rasul. "Social incentives in the workplace." The Review of Economic Studies 77.2 (2010): 417-458.

[4] Bandiera, Oriana, Iwan Barankay, and Imran Rasul. "Team incentives: evidence from a firm level experiment." Journal of the European Economic Association 11.5 (2013): 1079-1114.

[5] Barankay, Iwan. "Rank incentives: Evidence from a randomized workplace experiment." Discussion papers (2012).

[6] Benabou, Roland, and Jean Tirole. "Intrinsic and extrinsic motivation." The Review of Economic Studies 70.3 (2003): 489-520.

[7] Blader, Steven L., and Tom R. Tyler. "Testing and extending the group engagement model: linkages between social identity, procedural justice, economic outcomes, and extrarole behavior." Journal of Applied Psychology 94.2 (2009): 445.

[8] Blanes i Vidal, Jordi, and Mareike Nossol. "Tournaments without prizes: Evidence from personnel records." Management Science 57.10 (2011): 1721-1736.

[9] Bloom, Nicholas, and John Van Reenen. Measuring and explaining management practices across firms and countries. The Quarterly Journal of Economics (2007) 122 (4).

[10] Bloom, Nicholas, and John Van Reenen. "Why do management practices differ across firms and countries?." The Journal of Economic Perspectives (2010): 203-224.

[11] Branscombe, N. R., N. Ellemers, R. Spears, and E. J. Doosje. "The context and content of social identity threat." Social identity: Context, commitment, content (1999): 35-55.

[12] Brown, Rupert, and Dominic Abrams. "The effects of intergroup similarity and goal interdependence on intergroup attitudes and task performance." Journal of Experimental Social Psychology 22.1 (1986): 78-92.

[13] Brynjolfsson, Erik., and Paul. Milgrom. "Complementarity in Organizations". in (R. Gibbons and J. Roberts ed.) The Handbook of Organizational Economics.(2013).

[14] Bursztyn, Leonardo and Robert Jensen. "How Does Peer Pressure Affect Educational Investments?" Quarterly Journal of Economics, 130(3) (2015).

[15] Cowen, Tyler. Average is over: powering America beyond the age of the great stagnation. Penguin, 2013.

[16] Ellemers, Naomi, Russell Spears, and Bertjan Doosje. "Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility." Journal of Personality and Social Psychology 72.3 (1997): 617.

[17] Garcia, Stephen M., Avishalom Tor, and Richard Gonzalez. "Ranks and rivals: A theory of competition." Personality and Social Psychology Bulletin 32.7 (2006): 970-982.

[18] Garcia, Stephen M., and Avishalom Tor. "The n-effect more competitors, less competition." Psychological Science 20.7 (2009): 871-877.

[19] Gibbons, Robert and Rebecca Henderson, "What do Managers Do?" in (R. Gibbons and J. Roberts ed.) The Handbook of Organizational Economics, (2013).

[20] Gruenfeld, Deborah H., and Larissa Z. Tiedens. "Organizational preferences and their consequences." Handbook of social psychology (2010).

[21] Halevy, Nir, Eileen Y. Chou, Adam D. Galinsky, and J. Keith Murnighan. "When hierarchy wins evidence from the national basketball association." Social Psychological and Personality Science 3, no. 4 (2012): 398-406.

[22] Haran, Uriel, and Ilana Ritov. "Know who you're up against: Counterpart identifiability enhances competitive behavior." Journal of Experimental Social Psychology 54 (2014): 115-121.

[23] Haslam, S. Alexander. Psychology in organizations. Sage, 2004.

[24] Helper, Susan, and Rebecca Henderson. "Management Practices, Relational Contracts, and the Decline of General Motors." Journal of Economic Perspectives 28.1 (2014): 49-72.

[25] Holweg, Matthias. "The genealogy of lean production." Journal of operations management 25.2 (2007): 420-437.

[26] Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." American Economic Review 87.3 (1997): 291-313.

[27] Ichniowski, Casey, and Kathryn Shaw. "Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices." Journal of Economic Perspectives (2003): 155-180.

[28] Kandel, Eugene, and Edward P. Lazear. "Peer pressure and partnerships."Journal of political Economy (1992): 801-817.

[29] Kilduff, Gavin J., Hillary Anger Elfenbein, and Barry M. Staw. "The psychology of rivalry: A relationally dependent analysis of competition." Academy of Management Journal 53.5 (2010): 943-969.

[30] Koeth, Denise. Mandated EOBRs/ELDs. *Fleet Equipment Magazine*, July 2013.

[31] Lawler, Edward, Susan Mohrman and Gerald E. Ledford Jr. 1995. Creating High Performance Organizations: Practices and Results of Employee Involvement and TQM in Fortune 1000 Companies. San Francisco: Jossey-Bass.

[32] Lawler, Edward, Susan Mohrman, and George Benson, 2001. Organizing for High Performance: Employee Involvement, TQM, Reengineering, and Knowledge Management in the Fortune 1000. San Francisco: Jossey-Bass.

[33] Lawler, Edward and Susan Mohrman, 2003. Creating a Strategic Human Resources Organization: An Assessment of Trends and New Directions. Stanford University Press.

[34] Lazear, Edward P., and Sherwin Rosen. "Rank-Order Tournaments as Optimum Labor Contracts." The Journal of Political Economy 89.5 (1981): 841-864.

[35] Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. "Performance Pay and Wage Inequality*." The Quarterly journal of economics 124.1 (2009): 1-49.

[36] Levin, Jonathan (2003), Relational Incentive Contracts, The American Economic Review, Vol. 93, No. 3 (Jun., 2003), pp. 835-857.

[37] Levinthal, Daniel A. "Adaptation on rugged landscapes." Management science43.7 (1997): 934-950.

[38] MacLeod, W. Bentley and James M. Malcomson Econometrica, Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment Vol. 57, No. 2 (Mar., 1989), pp. 447-480

[39] Mael, Fred, and Blake E. Ashforth. "Alumni and their alma mater: A partial test of the reformulated model of organizational identification." Journal of organizational Behavior 13.2 (1992): 103-123.

[40] Milgrom, Paul, and John Roberts. "The economics of modern manufacturing: Technology, strategy, and organization." The American Economic Review(1990): 511-528.

[41] Milgrom, Paul, and John Roberts. "Complementarities and fit strategy, structure, and organizational change in manufacturing." Journal of accounting and economics 19.2 (1995): 179-208.

[42] Porter, Michael, and Nicolaj Siggelkow. "Contextuality within activity systems and sustainability of competitive advantage." The Academy of Management Perspectives 22.2 (2008): 34-56.

[43] Rivkin, Jan W. "Imitation of complex strategies." Management science 46.6 (2000): 824-844.

[44] Spears, Russell, Bertjan Doosje, and Naomi Ellemers. "Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification." Personality and Social Psychology Bulletin 23.5 (1997): 538-553.

[45] Stanne, Mary Beth, David W. Johnson, and Roger T. Johnson. "Does competition enhance or inhibit motor performance: a meta-analysis." Psychological bulletin 125.1 (1999): 133.

[46] Syverson, Chad (2011) What Determines Productivity? Journal of Economic Literature 2011, 49:2, 326–365

[47] Tajfel, Henri, and John C. Turner. "An integrative theory of intergroup conflict." The social psychology of intergroup relations 33.47 (1979): 74.

[48] Turner, John C. "Social categorization and the self-concept: A social cognitive theory of group behavior." Advances in group processes 2 (1985): 77-122.

[49] Womack, Jones, and D. Jones and D. Roos (1990). The Machine That Changed the World."

Figures

Figure 1: Impact of Rankings on Driver Performance

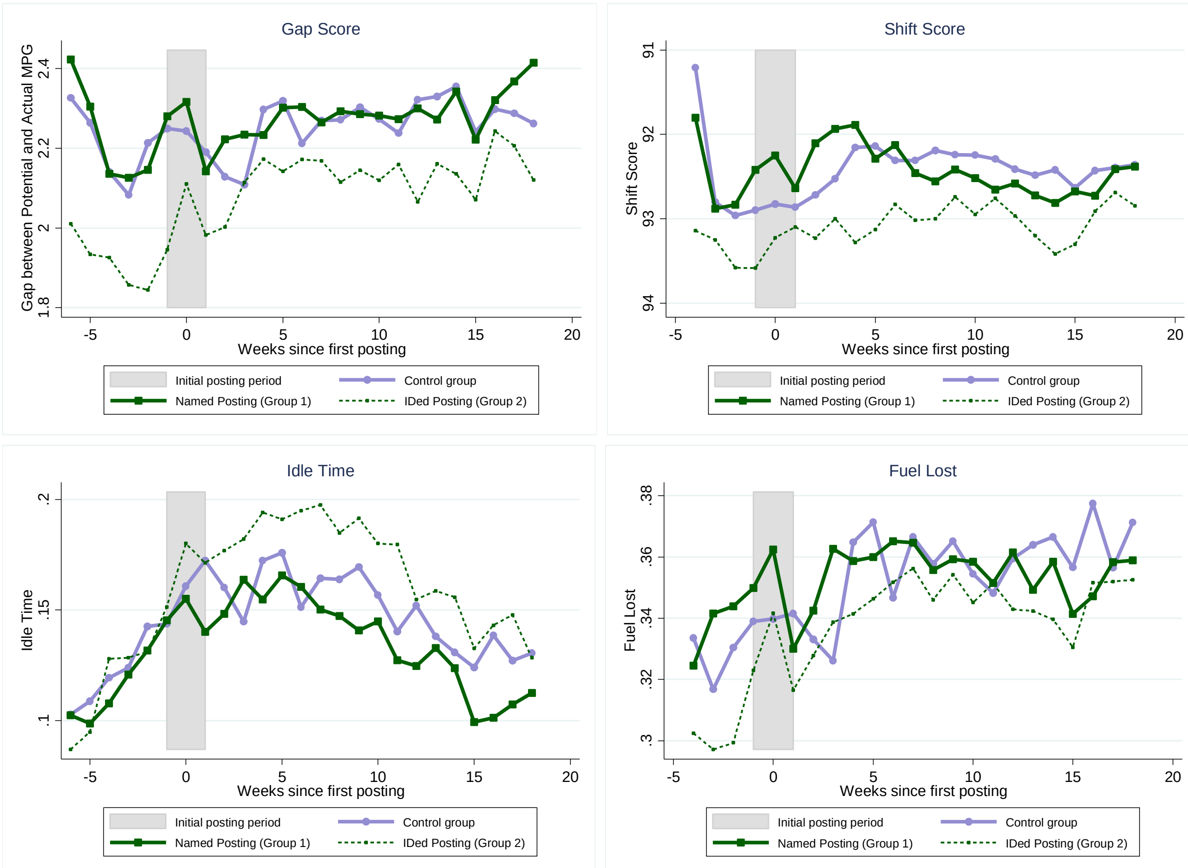


Figure 2: Impact of Rankings by Site Type

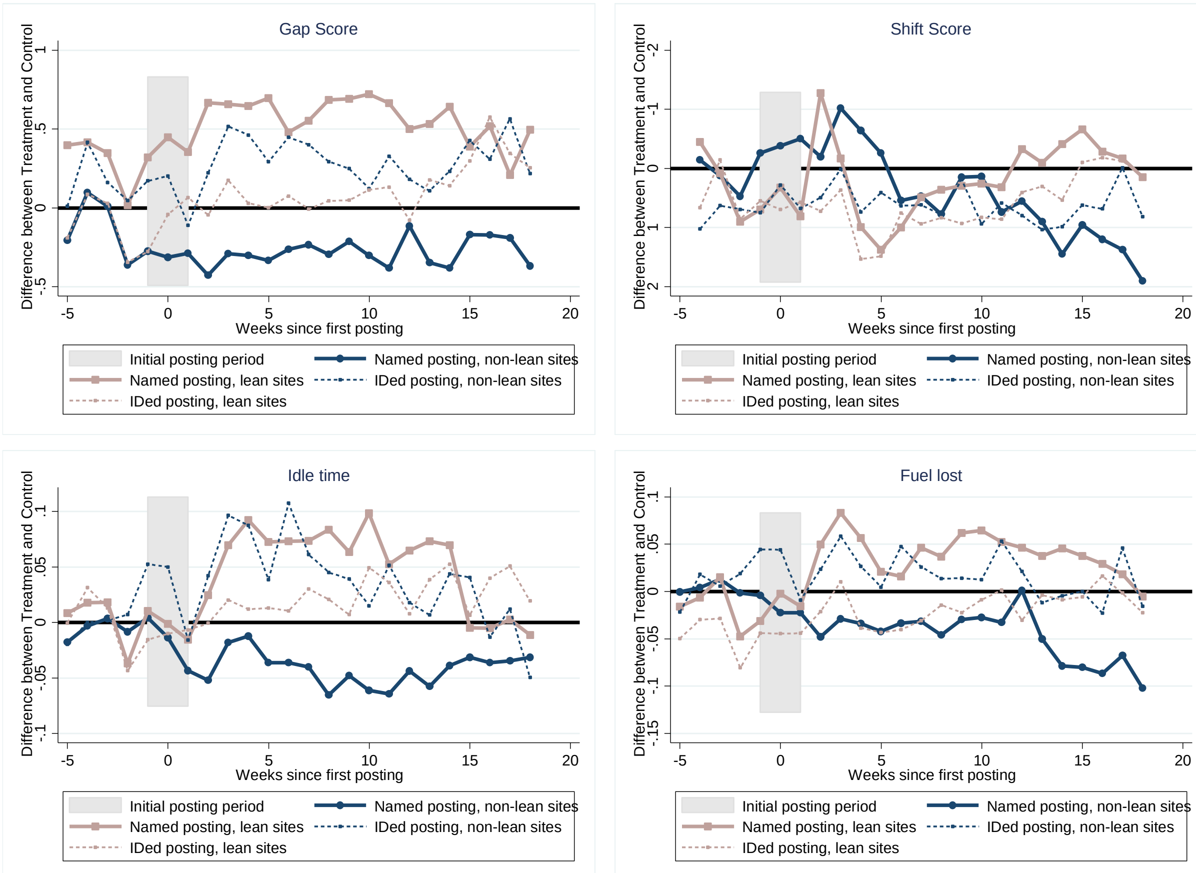
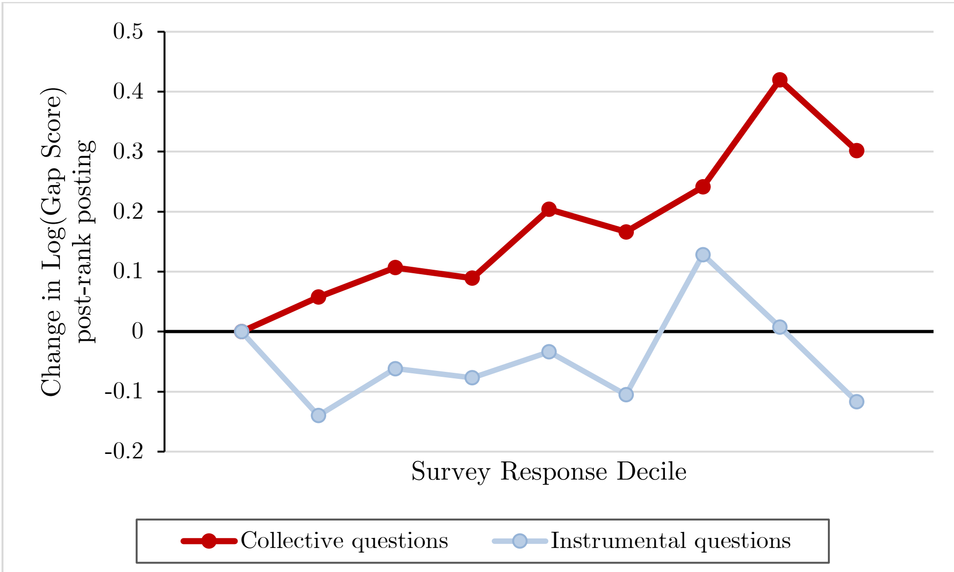
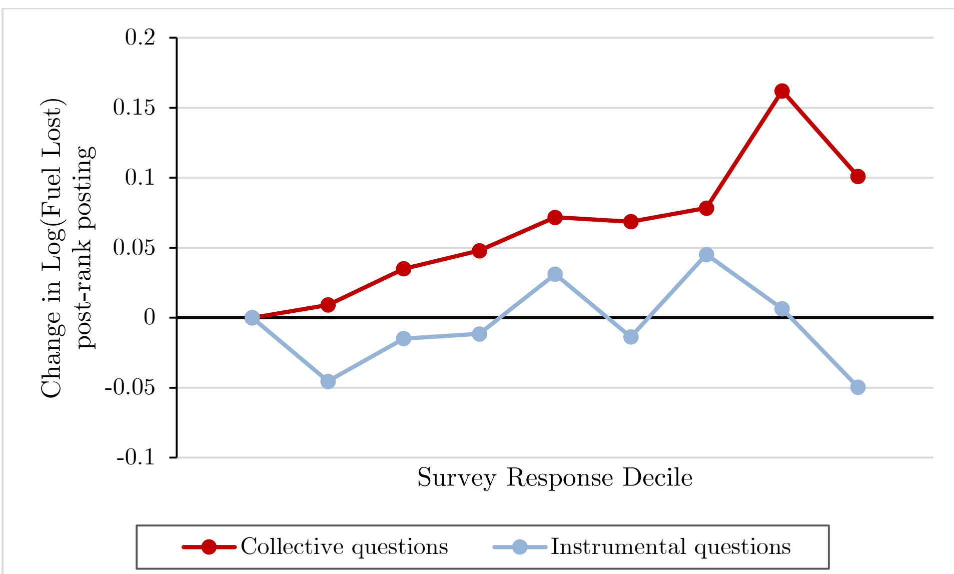


Figure 3: Rank posting response by engagement survey response





Tables

Table 1: Descriptive statistics

	All Sites				Sample				Control	Treatment 1 (names)	Diff	Treatment 2 (IDs)	Diff
	Mean	Med	Min	Max	Mean	Med	Min	Max	Mean	Mean	p-value	Mean	p-value
Site characteristics													
# sites	XXX*	n/a	n/a	n/a	142	n/a	n/a	n/a	47	50	n/a	45	n/a
Lean status	0.37	n/a	0	1	0.34	n/a	0	1	0.30	0.26	0.681	0.47	0.098
Tractors / site	30.96	22	5	151	24.71	21	5	87	25	25.32	0.924	23.73	0.664
Distance / trip	127.64	127.42	43.76	196.46	128.65	128.5	56.08	200.65	124.08	130.63	0.309	131.24	0.247
Pre-ranking driver performance (site mean)													
Miles per gallon	6.81	6.83	5.35	8.23	6.82	6.80	5.66	8.37	6.76	6.88	0.247	6.82	0.558
Gap score	2.19	2.1	0.57	6.99	2.10	1.94	0.78	6.43	2.18	2.14	0.787	1.98	0.310
Shift score	90.82	91.2	74.78	97.28	91.07	91.53	74.43	97.41	90.77	90.69	0.902	91.79	0.149
Excess idle time	0.12	0.1	0	0.72	0.13	0.11	0.02	0.72	0.12	0.12	0.838	0.14	0.429
Fuel lost	0.33	0.32	0.12	0.74	0.33	0.33	0.14	0.72	0.34	0.35	0.722	0.31	0.185

* The total number of sites in the firm has been masked for confidentiality purposes.

Table 2: Effect of rankings on all sites

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1	0.0061	0.0107	-0.6105	-0.6979*	0.0004	0.0007	-0.0007	0.0029
	(0.0213)	(0.0206)	(0.3902)	(0.3736)	(0.0076)	(0.0072)	(0.0077)	(0.0073)
Post* Posting group 2	0.0429	0.0490**	-0.3728	-0.4428	0.0138*	0.0124*	0.0118	0.0128*
	(0.0262)	(0.0221)	(0.3786)	(0.3622)	(0.0077)	(0.0074)	(0.0082)	(0.0076)
Post	0.0139	0.0061	0.1084	0.2289	0.0216***	0.0188***	0.0079	0.0026
	(0.0168)	(0.0149)	(0.3121)	(0.3016)	(0.0047)	(0.0046)	(0.0054)	(0.0051)
Posting group 1 (names)	-0.0030	-0.0096	0.1024	0.0324	-0.0048	-0.0116	0.0081	-0.0020
	(0.0445)	(0.0262)	(0.4995)	(0.3689)	(0.0088)	(0.0077)	(0.0118)	(0.0099)
Posting group 2 (IDs)	-0.0845*	-0.0147	0.8754*	0.0980	0.0008	0.0110	-0.0161	-0.0043
	(0.0477)	(0.0363)	(0.4861)	(0.4372)	(0.0085)	(0.0072)	(0.0129)	(0.0142)
Constant	0.9366***	1.7845***	92.5090***	90.4229***	0.1001***	0.2794***	0.2524***	-0.0842*
	(0.0327)	(0.1247)	(0.3352)	(1.4176)	(0.0056)	(0.0366)	(0.0080)	(0.0476)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	93913	93913	93913	93913	93913	93913	93913	93913
Adjusted R-squared	0.003	0.128	0.003	0.044	0.005	0.050	0.002	0.102

Table 3: Effect of rankings on lean and non-lean sites

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.1358***	0.1338***	-1.8625**	-1.9657***	0.0394***	0.0352***	0.0593***	0.0538***
	(0.0384)	(0.0383)	(0.7702)	(0.7120)	(0.0136)	(0.0131)	(0.0144)	(0.0127)
Post*Posting group 2*Lean	0.0251	0.0329	-0.4615	-0.6557	-0.0080	-0.0073	0.0093	0.0131
	(0.0522)	(0.0462)	(0.7701)	(0.7176)	(0.0159)	(0.0152)	(0.0168)	(0.0161)
Post*Posting group 1	-0.0423*	-0.0375	0.0904	0.0366	-0.0130	-0.0113	-0.0217***	-0.0164**
	(0.0251)	(0.0232)	(0.4018)	(0.3800)	(0.0084)	(0.0080)	(0.0077)	(0.0080)
Post*Posting group 2	0.0329	0.0333	-0.2777	-0.2295	0.0178	0.0160	0.0079	0.0064
	(0.0392)	(0.0336)	(0.3835)	(0.3777)	(0.0134)	(0.0128)	(0.0116)	(0.0117)
Post*Lean	-0.0303	-0.0269	0.9783	1.0118*	0.0034	0.0027	-0.0098	-0.0090
	(0.0326)	(0.0295)	(0.6547)	(0.6043)	(0.0091)	(0.0090)	(0.0111)	(0.0100)
Posting group 1*Lean	0.1434*	0.0415	0.5095	0.7454	0.0066	-0.0166	0.0149	0.0066
	(0.0864)	(0.0658)	(0.9945)	(0.7343)	(0.0180)	(0.0163)	(0.0242)	(0.0241)
Posting group 2*Lean	0.0549	0.0814	-0.7479	-0.3329	-0.0002	0.0141	-0.0022	0.0294
	(0.0895)	(0.0682)	(0.9417)	(0.7676)	(0.0165)	(0.0125)	(0.0250)	(0.0266)
Post	0.0263	0.0180	-0.2856	-0.1838	0.0203***	0.0178***	0.0118*	0.0066
	(0.0222)	(0.0186)	(0.2532)	(0.2425)	(0.0064)	(0.0063)	(0.0065)	(0.0065)
Lean	-0.0937	-0.0814*	0.1299	0.2731	-0.0014	-0.0123	-0.0139	-0.0264
	(0.0590)	(0.0445)	(0.6939)	(0.5931)	(0.0106)	(0.0093)	(0.0160)	(0.0169)
Posting group 1 (names)	-0.0578	-0.0296	-0.0684	-0.1717	-0.0072	-0.0071	0.0022	-0.0062
	(0.0575)	(0.0337)	(0.6552)	(0.4627)	(0.0116)	(0.0102)	(0.0146)	(0.0131)
Posting group 2 (IDs)	-0.0996	-0.0503	1.2832**	0.2648	0.0012	0.0049	-0.0124	-0.0174
	(0.0659)	(0.0426)	(0.5405)	(0.5338)	(0.0121)	(0.0104)	(0.0162)	(0.0166)
Constant	0.9741***	1.7811***	92.4570***	90.7710***	0.1007***	0.2767***	0.2579***	-0.0854*
	(0.0453)	(0.1292)	(0.4212)	(1.4685)	(0.0081)	(0.0371)	(0.0098)	(0.0488)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	93913	93913	93913	93913	93913	93913	93913	93913
Adjusted R-squared	0.009	0.130	0.004	0.045	0.007	0.051	0.004	0.104

Notes: SE clustered by SIC, window (5,30), controls include XYZ, winsorization

Table 4: Analysis using date and driver fixed effects

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.1026***	0.1067***	-1.9433***	-1.9459***	0.0279**	0.0254*	0.0501***	0.0441***
	(0.0359)	(0.0355)	(0.6079)	(0.6041)	(0.0140)	(0.0132)	(0.0141)	(0.0124)
Post*Posting group 2*Lean	0.0544	0.0561	-1.2449**	-1.2635**	-0.0087	-0.0069	0.0169	0.0192
	(0.0396)	(0.0389)	(0.5879)	(0.5845)	(0.0157)	(0.0152)	(0.0143)	(0.0136)
Post*Posting group 1	-0.0424*	-0.0452**	0.3998	0.3974	-0.0133	-0.0109	-0.0231***	-0.0180**
	(0.0229)	(0.0221)	(0.3492)	(0.3484)	(0.0087)	(0.0084)	(0.0079)	(0.0077)
Post*Posting group 2	0.0112	0.0115	0.0615	0.0674	0.0143	0.0131	0.0042	0.0021
	(0.0304)	(0.0292)	(0.3261)	(0.3252)	(0.0127)	(0.0126)	(0.0099)	(0.0104)
Post*Lean	-0.0368	-0.0415	1.2533**	1.2798***	0.0062	0.0052	-0.0130	-0.0126
	(0.0272)	(0.0271)	(0.4947)	(0.4896)	(0.0094)	(0.0090)	(0.0105)	(0.0092)
Posting group 1*Lean								
Posting group 2*Lean								
Post	-0.0049	-0.0027	-0.8095***	-0.8222***	-0.0069	-0.0064	-0.0066	-0.0067
	(0.0216)	(0.0208)	(0.2753)	(0.2741)	(0.0081)	(0.0080)	(0.0082)	(0.0080)
Lean								
Posting group 1 (names)								
Posting group 2 (IDs)								
Constant	0.9495***	1.8575***	92.9190***	88.6765***	0.1260***	0.1821***	0.2564***	-0.0576***
	(0.0119)	(0.0484)	(0.1367)	(0.5761)	(0.0057)	(0.0131)	(0.0052)	(0.0195)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	93913	93913	93913	93913	93913	93913	93913	93913
Adjusted R-squared	0.546	0.559	0.601	0.602	0.285	0.298	0.459	0.509

Table 5: Comparison of lean and non-lean sites

	Full sample			Matched sample		
	Non-lean Mean	Lean Mean	Diff p-value	Non-lean Mean	Lean Mean	Diff *p*-value
# sites	94	48	n/a	39	39	n/a
Tractors / site	20.35	33.25	0.000	25.95	27.51	0.581
Distance / trip	128.04	127.53	0.609	128.04	127.53	0.937
Miles per gallon	6.90	6.72	0.039	6.76	6.71	0.602
Gap score	2.14	2.04	0.537	2.00	2.03	0.838
Shift score	90.35	91.55	0.076	91.62	91.66	0.950
Excess idle time	0.12	0.13	0.781	0.12	0.13	0.815
Fuel lost	0.34	0.33	0.473	0.32	0.33	0.753
Eastern region	0.27	0.39	0.155	0.37	0.38	0.865
Central region	0.41	0.37	0.626	0.44	0.38	0.626
Western region	0.32	0.24	0.357	0.20	0.23	0.701
Control group	0.35	0.29	0.480	0.39	0.27	0.245
Posting group 1	0.39	0.27	0.149	0.32	0.24	0.467
Posting group 2	0.26	0.44	0.027	0.29	0.49	0.072

Table 6: Matched analysis

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.1485***	0.1510***	0.3944	0.0457	0.0523***	0.0522***	0.0585***	0.0582***
	(0.0493)	(0.0433)	(0.7437)	(0.7280)	(0.0172)	(0.0157)	(0.0174)	(0.0153)
Post*Posting group 2*Lean	0.0034	0.0254	0.7460	0.4590	-0.0203	-0.0151	0.0067	0.0084
	(0.0626)	(0.0532)	(0.6049)	(0.5673)	(0.0192)	(0.0181)	(0.0196)	(0.0193)
Post*Posting group 1	-0.0623*	-0.0496	-0.7339	-0.6517	-0.0214*	-0.0213*	-0.0238**	-0.0197*
	(0.0343)	(0.0327)	(0.5263)	(0.5547)	(0.0123)	(0.0115)	(0.0110)	(0.0112)
Post*Posting group 2	0.0350	0.0387	-0.3744	-0.3583	0.0298*	0.0258	0.0073	0.0094
	(0.0503)	(0.0443)	(0.3742)	(0.4068)	(0.0173)	(0.0166)	(0.0149)	(0.0153)
Post*Lean	-0.0435	-0.0425	-0.3965	-0.2911	-0.0017	-0.0050	-0.0139	-0.0126
	(0.0408)	(0.0316)	(0.4778)	(0.4521)	(0.0113)	(0.0106)	(0.0126)	(0.0122)
Posting group 1*Lean	0.0582	0.0384	0.3169	-0.8576	-0.0011	0.0015	-0.0131	0.0155
	(0.1096)	(0.0934)	(1.0673)	(0.9340)	(0.0253)	(0.0254)	(0.0268)	(0.0328)
Posting group 2*Lean	-0.0848	0.0842	-0.2996	-1.4198	-0.0109	0.0208	-0.0400	0.0379
	(0.1134)	(0.0825)	(1.0487)	(0.8966)	(0.0202)	(0.0173)	(0.0295)	(0.0306)
Post	0.0496*	0.0347	-0.1958	-0.0739	0.0245**	0.0216**	0.0173**	0.0113
	(0.0296)	(0.0250)	(0.2566)	(0.3057)	(0.0095)	(0.0094)	(0.0082)	(0.0086)
Lean	0.0135	-0.0331	-0.3308	0.7914	0.0022	-0.0156	0.0198	-0.0179
	(0.0839)	(0.0649)	(0.7787)	(0.7736)	(0.0138)	(0.0139)	(0.0189)	(0.0246)
Posting group 1 (names)	-0.0242	0.0064	0.0517	0.3504	-0.0003	-0.0074	0.0022	0.0078
	(0.0712)	(0.0628)	(0.7174)	(0.6182)	(0.0174)	(0.0192)	(0.0177)	(0.0240)
Posting group 2 (IDs)	0.0396	-0.0702	0.7376	0.8761	0.0094	0.0044	0.0105	-0.0263
	(0.0813)	(0.0680)	(0.6374)	(0.8571)	(0.0150)	(0.0142)	(0.0207)	(0.0261)
Constant	0.8846***	1.7351***	93.3009***	89.4941***	0.0973***	0.2914***	0.2427***	-0.0876
	(0.0574)	(0.1485)	(0.4419)	(1.6906)	(0.0100)	(0.0356)	(0.0126)	(0.0538)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	60002	60002	60002	60002	60002	60002	60002	60002
Adjusted R-squared	0.007	0.127	0.003	0.050	0.009	0.056	0.004	0.102

Table 7: Matched analysis with date and driver fixed effects

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.1088***	0.1092***	-0.1514	-0.1495	0.0360**	0.0350**	0.0453***	0.0435***
	(0.0392)	(0.0380)	(0.6861)	(0.6841)	(0.0172)	(0.0161)	(0.0152)	(0.0135)
Post*Posting group 2*Lean	0.0118	0.0214	-0.0348	-0.0812	-0.0267	-0.0234	0.0031	0.0043
	(0.0436)	(0.0426)	(0.5283)	(0.5263)	(0.0179)	(0.0175)	(0.0163)	(0.0158)
Post*Posting group 1	-0.0525*	-0.0514*	-0.4054	-0.4153	-0.0215*	-0.0196*	-0.0231**	-0.0204*
	(0.0308)	(0.0303)	(0.4953)	(0.4945)	(0.0114)	(0.0110)	(0.0114)	(0.0107)
Post*Posting group 2	0.0317	0.0274	-0.1981	-0.1755	0.0280*	0.0259*	0.0096	0.0080
	(0.0397)	(0.0388)	(0.3252)	(0.3275)	(0.0152)	(0.0152)	(0.0137)	(0.0141)
Post*Lean	-0.0299	-0.0355	0.0595	0.0869	0.0026	0.0006	-0.0078	-0.0086
	(0.0279)	(0.0268)	(0.4627)	(0.4589)	(0.0112)	(0.0109)	(0.0110)	(0.0102)
Posting group 1*Lean								
Posting group 2*Lean								
Post	0.0065	0.0106	-0.2165	-0.2353	-0.0070	-0.0059	-0.0018	-0.0019
	(0.0269)	(0.0259)	(0.3157)	(0.3164)	(0.0113)	(0.0110)	(0.0105)	(0.0102)
Lean								
Posting group 1 (names)								
Posting group 2 (IDs)								
Constant	0.9295***	1.8671***	93.4125***	89.6393***	0.1297***	0.2003***	0.2558***	-0.0455*
	(0.0161)	(0.0640)	(0.1689)	(0.6491)	(0.0073)	(0.0155)	(0.0072)	(0.0250)
Controls	N	Y	N	Y	N	Y	N	Y
Date and Driver Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Observations	60002	60002	60002	60002	60002	60002	60002	60002
Adjusted R-squared	0.559	0.572	0.620	0.621	0.289	0.303	0.470	0.521

Table 8: Effect on variance

Coefficient of variation:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	-0.0427*	-0.0436**	0.0044	0.0044	-0.2466***	-0.2466***	-0.0184	-0.0182
	(0.0239)	(0.0205)	(0.0053)	(0.0047)	(0.0787)	(0.0714)	(0.0354)	(0.0282)
Post*Posting group 2*Lean	0.0042	-0.0102	0.0003	0.0001	-0.0484	-0.0945	0.0265	0.0025
	(0.0241)	(0.0214)	(0.0050)	(0.0043)	(0.0776)	(0.0663)	(0.0339)	(0.0289)
Post*Posting group 1	0.0389***	0.0325***	-0.0019	-0.0015	0.1335***	0.1124***	0.0311*	0.0251
	(0.0129)	(0.0117)	(0.0029)	(0.0027)	(0.0421)	(0.0365)	(0.0177)	(0.0158)
Post*Posting group 2	-0.0169	-0.0097	-0.0011	-0.0012	-0.0171	0.0161	-0.0420*	-0.0286
	(0.0166)	(0.0145)	(0.0029)	(0.0027)	(0.0516)	(0.0413)	(0.0231)	(0.0196)
Post*Lean	-0.0059	-0.0010	-0.0032	-0.0027	0.0867	0.1178**	-0.0037	0.0026
	(0.0155)	(0.0142)	(0.0036)	(0.0031)	(0.0543)	(0.0479)	(0.0230)	(0.0197)
Posting group 1*Lean	-0.0034	-0.0024	0.0032	0.0056	-0.0029	0.0073	0.0038	-0.0208
	(0.0171)	(0.0155)	(0.0036)	(0.0034)	(0.0591)	(0.0556)	(0.0249)	(0.0210)
Posting group 2*Lean	-0.0573***	-0.0158	0.0076**	0.0146***	-0.1694***	-0.1565***	-0.0463*	0.0145
	(0.0173)	(0.0174)	(0.0034)	(0.0033)	(0.0581)	(0.0550)	(0.0244)	(0.0235)
Post	-0.0253***	-0.0199**	0.0070***	0.0064***	-0.2051***	-0.1862***	-0.0180	-0.0154
	(0.0087)	(0.0079)	(0.0018)	(0.0017)	(0.0289)	(0.0252)	(0.0120)	(0.0108)
Lean	0.0493***	0.0175	-0.0023	-0.0105***	0.1039**	0.0505	0.0447***	0.0120
	(0.0111)	(0.0120)	(0.0026)	(0.0025)	(0.0405)	(0.0423)	(0.0162)	(0.0164)
Posting group 1 (names)	0.0149	0.0067	0.0040**	0.0024	-0.0476	-0.0385	-0.0044	0.0066
	(0.0096)	(0.0091)	(0.0020)	(0.0020)	(0.0317)	(0.0293)	(0.0130)	(0.0122)
Posting group 2 (IDs)	0.0729***	0.0326***	-0.0091***	-0.0078***	0.0732*	0.0084	0.0463***	0.0019
	(0.0122)	(0.0121)	(0.0020)	(0.0022)	(0.0393)	(0.0351)	(0.0172)	(0.0163)
Constant	0.5712***	0.3053***	0.0733***	0.1375***	1.6574***	-0.0409	0.7965***	1.1360***
	(0.0066)	(0.0695)	(0.0012)	(0.0152)	(0.0224)	(0.1894)	(0.0086)	(0.0944)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	5343	5343	5343	5343	5285	5285	5343	5343
Adjusted R-squared	0.024	0.240	0.022	0.192	0.029	0.284	0.007	0.280

Table 9: Lean and employee engagement

Dependent variable:	Collective questions		Instrumental questions	
	(1)	(2)	(3)	(4)
Lean	0.2268	0.2735**	0.2954*	0.3314**
	(0.1606)	(0.1234)	(0.1753)	(0.1607)
Constant	3.3001***	3.2945***	3.2542***	2.7641***
	(0.1681)	(0.4946)	(0.1697)	(0.4358)
Demographic Controls	No	Yes	No	Yes
Observations	561	561	564	564
Adjusted R-squared	0.026	0.127	0.016	0.075

Table 10: Matched analysis of Lean and employee engagement

Dependent variable:	Collective questions		Instrumental questions	
	(1)	(2)	(3)	(4)
Lean	0.1731	0.2736**	0.1730	0.2353
	(0.1577)	(0.1298)	(0.1978)	(0.1965)
Constant	3.3550***	3.3327***	3.3138***	2.9711***
	(0.2183)	(0.3734)	(0.1881)	(0.6146)
Demographic Controls	No	Yes	No	Yes
Observations	396	396	399	399
Adjusted R-squared	0.029	0.128	0.009	0.056

Table 11: Effect of ranking and engagement on driver performance

Dependent variable:	Log(Gap Score)			
Category:	Collective questions		Instrumental questions	
	(1)	(2)	(3)	(4)
Post*Posting group 1*[Category]	0.1151***	0.1182**	0.0407	0.0493
	(0.0423)	(0.0526)	(0.0368)	(0.0488)
Post*Posting group 2*[Category]	0.0322	0.0042	-0.0243	-0.0517
	(0.0629)	(0.0815)	(0.0574)	(0.1012)
Post*Posting group 1	-0.4082***	-0.4440***	-0.1741	-0.2261
	(0.1273)	(0.1531)	(0.1157)	(0.1516)
Post*Posting group 2	-0.0511	-0.0500	0.1370	0.1377
	(0.2269)	(0.3196)	(0.2195)	(0.4129)
Post*[Category]	-0.0628*	-0.0554	-0.0196	-0.0248
	(0.0339)	(0.0455)	(0.0264)	(0.0391)
Posting group 1*[Category]	0.1151*	0.1384**	0.0876	0.1237
	(0.0600)	(0.0656)	(0.0671)	(0.0749)
Posting group 2*[Category]	-0.0578	0.0241	-0.0544	-0.0516
	(0.0677)	(0.0816)	(0.0596)	(0.0835)
Post	0.2570**	0.2497*	0.1246	0.1599
	(0.1027)	(0.1316)	(0.0914)	(0.1224)
[Category]	-0.0309	-0.0515	-0.0404	-0.0416
	(0.0397)	(0.0405)	(0.0461)	(0.0485)
Posting group 1 (Names)	-0.3260	-0.3650	-0.2278	-0.3003
	(0.2179)	(0.2294)	(0.2523)	(0.2664)
Posting group 2 (IDs)	0.2767	0.0437	0.2623	0.3041
	(0.2145)	(0.2581)	(0.1981)	(0.2679)
Constant	2.1358***	2.3431***	2.0377***	2.1163***
	(0.3214)	(0.3606)	(0.2985)	(0.3487)
Sample	Full	Matched	Full	Matched
Observations	35187	26065	35385	26263
Adjusted R-squared	0.106	0.117	0.095	0.112

Appendix Figures and Tables

Figure A1: Sample Rank Posting

Name	Total Distance (mi)	Total Fuel Consumption (gal)	Current MPG	Fuel Lost (gal)	Potential MPG	Potential Savings (%)	Shift Score (%)	Potential MPG	Excessive Idle Time	Excessive Idle Fuel Loss (gal)	Progressive Shifting Fuel Loss (gal)	Excessive Speed Fuel Loss	Idle Shutdowns	Max RPM
[illegible]	131.33	23.92	5.49	1.14	5.77	4.79	84	5.77	0 hr 34 m	0.28	0.87	0	1	2247
[illegible]	586.49	83.63	7.01	1.01	7.1	1.21	98	7.1	0 hr 42 m	0.21	0.74	0	0	1835
[illegible]	434.88	70.9	6.13	1.99	6.31	2.81	88	6.31	0 hr 18 m	0.17	1.79	0.04	0	1931
[illegible]	212.77	31.05	6.85	1.91	7.3	6.14	83	7.3	0 hr 14 m	0.13	1.77	0	0	2235
[illegible]	240.94	41.17	5.85	1.55	6.08	3.78	100	6.08	3 hr 13 m	1.54	0.01	0	0	1556
[illegible]	443.01	69.61	6.36	5.26	6.89	7.56	67	6.89	0 hr 12 m	0.16	4.92	0	0	2520
[illegible]	200.03	28.92	6.92	2.19	7.48	7.58	72	7.48	0 hr 4 mi	0.05	2.08	0.02	0	2226
[illegible]	2334.15	462.42	5.05	2.45	5.07	[illegible]	73	5.07	0 hr 27 m	0.18	1.78	0.17	0	2258
[illegible]	494.75	71.55	6.91	1.45	[illegible]	[illegible]	92	7.06	0 hr 3 mi	0.01	1.41	0.01	0	1799
[illegible]	836.66	119.79	6.98	2.87	[illegible]	[illegible]	[illegible]	7.16	4 hr 7 mi	1.45	1.14	0.06	0	2229
[illegible]	329.34	42.37	7.77	1.9	[illegible]	[illegible]	[illegible]	8.14	0 hr 53 m	1.5	0.38	0.02	0	1781
[illegible]	531.97	89.11	5.97	8.96	[illegible]	[illegible]	[illegible]	6.64	0 hr 12 m	0.15	8.1	0.02	1	3373
[illegible]	220.04	37.7	5.84	1.35	[illegible]	[illegible]	[illegible]	6.02	2 hr 4 mi	0.53	0.62	0	0	2171
[illegible]	290.93	48.12	6.05	1.7	[illegible]	[illegible]	[illegible]	6.27	3 hr 53 m	1.59	0.1	0	0	1679
[illegible]	316.08	46.04	6.87	0.57	6.95	[illegible]	95	6.95	0 hr 22 m	0.11	0.46	0	0	1857
[illegible]	611.71	102.35	5.98	1.83	6.09	1.78	96	6.09	0 hr 12 m	0.06	1.56	0.03	0	1882
[illegible]	160.58	23.93	6.71	1.02	7.01	4.28	69	7.01	0 hr 7 mi	0.05	0.96	0	0	2534
[illegible]	300.02	54.69	5.49	2.27	5.72	4.16	81	5.72	0 hr 22 m	0.14	2.13	0	0	2155
[illegible]	233.75	37.15	6.29	2.05	6.66	5.51	83	6.66	1 hr 14 m	0.35	1.69	0	0	1994
[illegible]	293.76	43.37	6.77	0.35	6.83	0.81	99	6.83	0 hr 10 m	0.11	0.23	0.01	0	1816
[illegible]	1761.11	264.44	6.66	6.9	6.84	2.61	94	6.84	0 hr 27 m	0.3	5.58	0.39	0	1906
[illegible]	1724.74	299.13	5.77	5	5.86	1.67	59	5.86	4 hr 8 mi	1.56	0.97	2.52	0	2306
[illegible]	370.28	56.31	6.58	0.98	6.69	1.74	100	6.69	4 hr 38 m	0.93	0.05	0	1	1799
[illegible]	100.83	19.14	5.27	0.27	5.35	1.43	43	5.35	0 hr 12 m	0.07	0.2	0	0	2289
[illegible]	239.08	37.05	6.45	1.21	6.67	3.26	83	6.67	0 hr 14 m	0.07	1.14	0	1	1982
[illegible]	20.42	2.68	7.61	0.08	7.84	2.98	58	7.84	0 hr 1 mi	0.02	0.06	0	0	2304
[illegible]	619.79	79.39	7.81	1.06	7.91	1.33	95	7.91	0 hr 0 mi	0.01	0.86	0.04	0	2192
[illegible]	310.84	47	6.61	0.83	6.73	1.76	89	6.73	0 hr 1 mi	0.02	0.81	0	0	1862
[illegible]	306.04	37.34	8.17	0.86	8.36	2.3	88	8.36	0 hr 1 mi	0.01	0.85	0	0	1890
[illegible]	211.26	38.62	5.47	1.44	5.68	3.72	87	5.68	0 hr 0 mi	0	1.44	0	0	1864
[illegible]	193.17	26.96	7.16	0.17	7.21	0.62	0	7.21	0 hr 2 mi	0.02	0.13	0.01	0	1865
[illegible]	386.71	61.49	6.29	5.74	6.94	9.33	74	6.94	7 hr 34 m	3.82	1.92	0	0	2363
[illegible]	145.47	19.15	7.6	1.01	8.02	5.29	80	8.02	0 hr 14 m	0.17	0.83	0	0	2225
[illegible]	240.12	41.23	5.82	0.95	5.96	2.3	94	5.96	0 hr 6 mi	0.06	0.89	0.01	0	1790
[illegible]	183.93	27.59	6.67	1.55	7.06	5.63	84	7.06	1 hr 47 m	0.33	1.21	0	0	2049
[illegible]	216.09	35.49	6.09	2.21	6.49	6.22	72	6.49	1 hr 10 m	0.31	1.9	0	0	1887
[illegible]	453.12	65.48	6.92	1.31	7.06	1.99	88	7.06	0 hr 40 m	0.3	0.99	0.01	0	1864
[illegible]	0.46	0.18	2.6	0	2.61	0.7	99	2.61	0 hr 0 mi	0	0	0	0	1481
[illegible]	245.95	39.15	6.23	3.9	6.92	9.97	53	6.92	0 hr 10 m	0.09	3.49	0	0	2733
[illegible]	182.83	26.74	6.84	2.36	7.5	8.81	87	7.5	1 hr 58 m	0.77	1.58	0	0	1836
[illegible]	61.45	8.54	7.2	0.24	7.41	2.79	91	7.41	0 hr 1 mi	0.01	0.23	0	0	1854
[illegible]	642.67	87.5	7.34	0.88	7.42	1	95	7.42	0 hr 8 mi	0.05	0.81	0	1	1999
[illegible]	144.49	21.78	6.64	0.87	6.91	4.01	83	6.91	0 hr 0 mi	0	0.85	0	0	2671
[illegible]	6.46	0.77	8.39	0	8.42	0.34	98	8.42	0 hr 0 mi	0	0	0	0	2237
[illegible]	118.53	21.68	5.47	1.18	5.78	5.44	48	5.78	0 hr 35 m	0.12	0.86	0	0	4916
[illegible]	187.81	24.65	7.62	0.29	7.71	1.18	98	7.71	0 hr 16 m	0.11	0.18	0	0	1754
[illegible]	390.19	54.05	7.22	0.75	7.32	1.4	98	7.32	0 hr 6 mi	0.06	0.52	0.09	0	1791
[illegible]	517.46	87.72	5.9	5.3	6.28	6.05	77	6.28	0 hr 15 m	0.09	4.95	0.05	1	2213
	19181.47	3059.01	6.27	90.98	6.46	2.97	84	6.46	1 day 20	18.26	66.01	3.53	6	2124

Figure A2: Lean Evaluation Criteria

Safety	Employee s have a formal avenue to openly voice, share, and regularly address safety concerns at the facility
	Safety concerns are addressed in a timely manner by a cross-functional, integrated team of employees, supervision, and management.
Safety and leadership	What level of leader is involved in the safety journey?
	What organizational levels originated, supported, and have advocated the lean implementation initiative in the facility?
Power distance	Management availability to team members. Do employees feel that management is approachable?
	What percentage of the day do Supervisors spend on the Dock, during normal working hours?
	What percentage of the day do Managers spend on the Dock, during normal working hours?
Employee recognition	Individuals who meet, exceed, or achieve objectives are recognized on a regular basis through an employee recognition program?
	Groups who meet, exceed, or achieve objectives are recognized on a regular basis through a group recognition program?
Management style	Feedback and concerns are encouraged and included before making changes and taking actions.
	Employees, Supervisors, and Managers are encouraged/empowered to try improvement ideas, using innovation and creativity to enrich job responsibilities.
	The organizational level involved in determining and leading facility, function, and CIR Goals.
Teamwork and empowerment	Daily work activities are organized into team functions.
	SME s are utilized as initial point of contact for problem-solving, resolution, and employee directing activities.
	Problem-Solving activities are organized into team functions.
	Employees are empowered, utilized, participate, initiate, and lead problem-solving activities autonomously, without significant management involvement.
Communication	There is an avenue for workers to openly share common concerns, issues, and problems regularly with other employees, supervisors, and management.
	Employee concerns and questions are addressed in a timely manner.
	Are there daily meetings with employees and supervision/management where the daily plans, performance, etc. are shared?

Figure A3: Samples from Interviews on Lean

Supervisor on how lean has affected his management style	"These guys will do anything for me, and they'll do absolutely nothing for other people. And I learned a lot of that from lean because lean has made me softer, it really has. I used to be hard as rock and now I feel like I'm a spongefiI still have that same pride but it's – my interaction with people is so much different, it's so much different. You're not treating them in a negative way or a negative manner and that's – I was hard as a rock in my numbers produced fiand if somebody didn't want to get on board with me on my team in all likelihood it probably wasn't going to be a very good day for that person. Now, it's with everybody being involved instead of just me running the show, it's totally different. Yes, are my numbers as good? Probably not, but you know what I'll take that. I firmly believe I'm a better supervisor today than what I was 6 months back."
Supervisor #2 on how lean has motivated drivers	"Since lean was introduced it was sort of like the door opening up. [Manager said] give it a chance, look at it and see what it can do. And I tell you it can produce productivity out of people that you thought would never produce. All it takes is a little bit of respect, little bit of understanding, show these guys that they're part of the operation."
Driver #1 on how lean has created community	"These guys now they get together, we got great relationships outside of the work environment. We've been to some of their homes. We do the activities outside of work. Even though I got Friday nights about once every month I sneak on down to Fridays and I buy them all the drink. It's just made us such a cohesive team it's incredible."
Driver #2 on how lean has increased teamwork	"I guess we haven't really been able to do too much yet – but I think the meetings and stuff have actually helped just getting people working together. So in the lean team, I think there's actually a good amount of camaraderie going on. So I think that's actually been good. Now some people I didn't really get along and stuff are working together."

Table A1: Sample construction

Sample construction	Driver-days	Sites
Total driver-days	1,137,192	275
- less early lean sites	(173,461)	(25)
- less late Q3/Q4 2013 lean sites	(130,679)	(36)
- less pre-11/25 rank posting dates	(416,593)	(72)
- less line haul routes	(76,989)	0
- less uncalibrated data	(8,781)	0
Sample	330,689	142
Sample within 5-30 window	93,913	142

Table A2: Descriptive statistics for matched sample

	Matched sample				Control	Treatment 1 (names)	Diff	Treatment 2 (IDs)	Diff
	Mean	Median	Min	Max	Mean	Mean	p-value	Mean	p-value
Site characteristics									
# sites	82	n/a	n/a	n/a	27	23	n/a	32	n/a
Lean status	0.50	n/a	0	1	0.41	0.43	0.849	0.63	0.099
Tractors / site	26.73	22.50	8	61	24.96	29.61	0.213	26.16	0.694
Distance / trip	127.79	126.34	62.08	198.69	128.26	127.66	0.945	127.47	0.918
Pre-ranking driver performance									
Miles per gallon	6.74	6.71	5.67	7.77	6.60	6.80	0.125	6.80	0.083
Gap score	2.01	1.96	0.83	5.26	2.03	2.02	0.993	1.99	0.885
Shift score	91.64	92.25	82.67	97.37	90.92	91.83	0.290	92.10	0.150
Excess idle time	0.12	0.11	0.04	0.68	0.11	0.12	0.800	0.14	0.364
Fuel lost	0.32	0.31	0.15	0.71	0.33	0.31	0.762	0.32	0.554

Table A3: Placebo test

	Actual MPG		Potential MPG	
	(1)	(2)	(3)	(4)
Post*Treatment group 1*Lean	-0.0363***	-0.0223**	0.0158	-0.0560
	(0.0106)	(0.0100)	(0.0731)	(0.0527)
Post*Treatment group 2*Lean	-0.0059	-0.0163	0.0981	0.0679
	(0.0141)	(0.0107)	(0.0777)	(0.0541)
Post*Treatment group 1	0.0083	0.0082	0.0539	0.0589*
	(0.0071)	(0.0062)	(0.0495)	(0.0339)
Post*Treatment group 2	-0.0109	-0.0024	-0.0408	-0.0353
	(0.0114)	(0.0082)	(0.0562)	(0.0348)
Post*Lean	0.0068	0.0081	-0.1213**	-0.0579
	(0.0088)	(0.0074)	(0.0481)	(0.0383)
Treatment group 1*Lean	-0.0465*		-0.4144*	
	(0.0239)		(0.2316)	
Treatment group 2*Lean	-0.0123		-0.3375	
	(0.0242)		(0.2204)	
Post	-0.0026	0.0056	-0.0798**	0.0289
	(0.0060)	(0.0059)	(0.0345)	(0.0373)
Lean	0.0238		0.1512	
	(0.0158)		(0.1694)	
Treatment group 1 (names)	0.0168		0.1899	
	(0.0153)		(0.1493)	
Treatment group 2 (IDs)	0.0235		0.0918	
	(0.0186)		(0.1432)	
Constant	0.0212	-0.0470***	6.8736***	6.7390***
	(0.0197)	(0.0112)	(0.1061)	(0.0264)
Date and Driver FE	N	Y	N	Y
Observations	93913	93913	93913	93913
Adjusted R-squared	0.973	0.985	0.159	0.577

Table A4: Placebo test on matched sample

	Actual MPG		Potential MPG	
	(1)	(2)	(3)	(4)
Post*Treatment group 1*Lean	-0.0357***	-0.0202**	0.0368	-0.0272
	(0.0127)	(0.0097)	(0.1032)	(0.0695)
Post*Treatment group 2*Lean	-0.0015	-0.0050	0.1273	0.1413**
	(0.0160)	(0.0115)	(0.0945)	(0.0608)
Post*Treatment group 1	0.0117	0.0096	-0.0094	0.0595
	(0.0092)	(0.0078)	(0.0780)	(0.0499)
Post*Treatment group 2	-0.0093	-0.0081	-0.1068	-0.0812*
	(0.0138)	(0.0105)	(0.0725)	(0.0420)
Post*Lean	0.0079	0.0035	-0.0976	-0.0846*
	(0.0098)	(0.0061)	(0.0665)	(0.0464)
Treatment group 1*Lean	-0.0322		-0.1821	
	(0.0314)		(0.2904)	
Treatment group 2*Lean	0.0209		-0.0944	
	(0.0307)		(0.2387)	
Post	-0.0075	0.0042	-0.0749	0.0498
	(0.0077)	(0.0061)	(0.0513)	(0.0489)
Lean	0.0039		0.0043	
	(0.0224)		(0.1728)	
Treatment group 1 (names)	0.0107		0.3404	
	(0.0200)		(0.2122)	
Treatment group 2 (IDs)	-0.0118		0.1995	
	(0.0236)		(0.1770)	
Constant	0.0252	-0.0518***	6.7012***	6.6435***
	(0.0248)	(0.0136)	(0.1284)	(0.0346)
Date and Driver FE	N	Y	N	Y
Observations	60002	60002	60002	60002
Adjusted R-squared	0.973	0.985	0.159	0.571

Table A5: Instrumental variables analysis

Dependent variable:	log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.2553**	0.2962**	-4.3968**	-4.7937**	0.0837**	0.0793**	0.1175**	0.1194**
	(0.0997)	(0.1283)	(2.2285)	(2.3630)	(0.0344)	(0.0369)	(0.0502)	(0.0465)
Post*Posting group 2*Lean	0.0261	0.0325	-0.6834	-0.9888	-0.0281	-0.0260	0.0131	0.0177
	(0.1291)	(0.1184)	(1.6693)	(1.5435)	(0.0459)	(0.0435)	(0.0403)	(0.0392)
Post*Posting group 1	-0.0755*	-0.0703*	0.1585	0.0624	-0.0229	-0.0206	-0.0382***	-0.0302**
	(0.0441)	(0.0397)	(0.6976)	(0.6735)	(0.0145)	(0.0139)	(0.0138)	(0.0137)
Post*Posting group 2	0.0859	0.0880	-0.7348	-0.6465	0.0451	0.0391	0.0204	0.0178
	(0.1074)	(0.0973)	(1.0684)	(0.9919)	(0.0433)	(0.0407)	(0.0318)	(0.0317)
Post*Lean	-0.0303	-0.0300	0.9783	1.0158*	0.0034	0.0033	-0.0098	-0.0101
	(0.0325)	(0.0295)	(0.6523)	(0.6034)	(0.0091)	(0.0091)	(0.0111)	(0.0101)
Posting group 1*Lean	0.3318	0.0873	1.3078	1.9110	0.0111	-0.0518	0.0422	0.0123
	(0.2096)	(0.1675)	(2.3074)	(2.0494)	(0.0421)	(0.0497)	(0.0570)	(0.0609)
Posting group 2*Lean	0.1667	0.2357	-2.2287	-1.2104	-0.0010	0.0309	0.0028	0.0817
	(0.1887)	(0.1840)	(2.2929)	(1.8750)	(0.0384)	(0.0349)	(0.0540)	(0.0701)
Post	0.0263	0.0191	-0.2856	-0.1774	0.0203***	0.0180***	0.0118*	0.0069
	(0.0221)	(0.0186)	(0.2523)	(0.2430)	(0.0064)	(0.0064)	(0.0065)	(0.0065)
Lean	-0.0937	-0.0900*	0.1299	0.3012	-0.0014	-0.0130	-0.0139	-0.0283
	(0.0587)	(0.0467)	(0.6914)	(0.6381)	(0.0106)	(0.0100)	(0.0159)	(0.0174)
Posting group 1 (names)	-0.1011	-0.0597	-0.1197	-0.3691	-0.0126	-0.0135	0.0039	-0.0136
	(0.0982)	(0.0674)	(1.1465)	(0.9449)	(0.0200)	(0.0210)	(0.0257)	(0.0262)
Posting group 2 (IDs)	-0.2553*	-0.1596	3.2883*	0.8652	0.0030	0.0029	-0.0317	-0.0540
	(0.1470)	(0.1035)	(1.7225)	(1.3275)	(0.0312)	(0.0298)	(0.0398)	(0.0387)
Constant	0.9741***	1.7826***	92.4570***	90.8150***	0.1007***	0.2884***	0.2579***	-0.0859*
	(0.0452)	(0.1297)	(0.4197)	(1.4618)	(0.0081)	(0.0371)	(0.0098)	(0.0484)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	93913	93913	93913	93913	93913	93913	93913	93913
Adjusted R-squared	0.011	0.126		0.041	0.002	0.045	0.002	0.101

Table A6: Instrumental variable analysis on matched sample

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.2880**	0.3091**	0.5184	-0.1058	0.1055***	0.1065**	0.1208**	0.1179***
	(0.1145)	(0.1201)	(1.6522)	(1.6767)	(0.0406)	(0.0414)	(0.0496)	(0.0427)
Post*Posting group 2*Lean	-0.1089	-0.0627	2.0143	1.7536	-0.1269	-0.1009	-0.0109	-0.0108
	(0.2861)	(0.2453)	(2.1440)	(2.2801)	(0.1382)	(0.1190)	(0.0778)	(0.0798)
Post*Posting group 1	-0.1185*	-0.0989*	-1.3615	-1.1202	-0.0399*	-0.0415*	-0.0441**	-0.0391**
	(0.0639)	(0.0553)	(1.2229)	(1.2327)	(0.0216)	(0.0214)	(0.0186)	(0.0188)
Post*Posting group 2	0.1779	0.1703	-1.4758	-1.5799	0.1425	0.1166	0.0379	0.0406
	(0.2785)	(0.2386)	(2.0199)	(2.1887)	(0.1375)	(0.1185)	(0.0745)	(0.0766)
Post*Lean	-0.0435	-0.0424	-0.3965	-0.2954	-0.0017	-0.0042	-0.0139	-0.0125
	(0.0405)	(0.0323)	(0.4749)	(0.4401)	(0.0112)	(0.0106)	(0.0125)	(0.0127)
Posting group 1*Lean	0.1259	0.1508	0.8134	-2.2238	-0.0030	0.0117	-0.0306	0.0677
	(0.2426)	(0.2461)	(2.2628)	(2.4720)	(0.0545)	(0.0591)	(0.0582)	(0.0892)
Posting group 2*Lean	-0.2500	0.4302	-2.3961	-5.5245	-0.0432	0.1004	-0.0973	0.1904
	(0.4089)	(0.3118)	(3.3864)	(3.3759)	(0.0767)	(0.0801)	(0.0986)	(0.1237)
Post	-0.0425	0.0035	0.0908	0.5247	-0.0006	-0.0259	0.0039	0.0126
	(0.1235)	(0.1413)	(1.2531)	(1.4532)	(0.0303)	(0.0413)	(0.0312)	(0.0559)
Lean	0.1700	-0.2977	3.1711	3.4086	0.0405	-0.0462	0.0451	-0.1169
	(0.3858)	(0.1901)	(3.0445)	(2.1764)	(0.0730)	(0.0607)	(0.0923)	(0.0755)
Posting group 1 (names)	0.0496*	0.0338	-0.1958	-0.0597	0.0245***	0.0214**	0.0173**	0.0109
	(0.0294)	(0.0254)	(0.2550)	(0.3086)	(0.0095)	(0.0094)	(0.0082)	(0.0089)
Posting group 2 (IDs)	0.0135	-0.0663	-0.3308	1.0935	0.0022	-0.0217	0.0198	-0.0346
	(0.0834)	(0.0801)	(0.7739)	(0.9326)	(0.0137)	(0.0167)	(0.0187)	(0.0309)
Constant	0.8846***	1.7607***	93.3009***	89.0525***	0.0973***	0.2958***	0.2427***	-0.0746
	(0.0570)	(0.1562)	(0.4391)	(1.6999)	(0.0100)	(0.0379)	(0.0125)	(0.0553)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	60002	60002	60002	60002	60002	60002	60002	60002
Adjusted R-squared	.	0.117	.	0.046	.	0.048	0.004	0.088

Table A7: Effect of rankings on lean and non-lean sites - long window

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.0966**	0.0963**	-1.6616	-1.4921	0.0259	0.0252	0.0465***	0.0373***
	(0.0441)	(0.0429)	(1.0370)	(1.0080)	(0.0163)	(0.0156)	(0.0156)	(0.0140)
Post*Posting group 2*Lean	0.0263	0.0297	-1.5130	-1.4498	-0.0036	0.0001	0.0130	0.0131
	(0.0486)	(0.0464)	(0.9732)	(0.9436)	(0.0183)	(0.0167)	(0.0185)	(0.0161)
Post*Posting group 1	-0.0414	-0.0526**	0.6129	0.6155	-0.0116	-0.0114	-0.0259***	-0.0201**
	(0.0270)	(0.0248)	(0.6369)	(0.6179)	(0.0082)	(0.0073)	(0.0093)	(0.0090)
Post*Posting group 2	0.0289	0.0295	-0.0243	-0.0978	0.0168	0.0141	0.0065	0.0069
	(0.0354)	(0.0325)	(0.7387)	(0.7135)	(0.0127)	(0.0117)	(0.0133)	(0.0113)
Post*Lean	-0.0280	-0.0332	0.3156	0.3168	0.0066	0.0023	-0.0124	-0.0115
	(0.0328)	(0.0317)	(0.7552)	(0.7297)	(0.0121)	(0.0111)	(0.0117)	(0.0103)
Posting group 1*Lean	0.1407	0.0377	0.3884	0.9014	0.0100	-0.0046	0.0105	0.0076
	(0.0870)	(0.0675)	(1.3413)	(1.1558)	(0.0168)	(0.0161)	(0.0227)	(0.0246)
Posting group 2*Lean	0.0653	0.0556	-0.1648	0.8462	-0.0000	0.0080	0.0007	0.0194
	(0.0852)	(0.0692)	(1.1586)	(1.0033)	(0.0145)	(0.0157)	(0.0230)	(0.0269)
Post	0.0471**	0.0532***	1.6888***	1.6883***	0.0170***	0.0189***	0.0229***	0.0198***
	(0.0196)	(0.0167)	(0.4700)	(0.4476)	(0.0060)	(0.0053)	(0.0065)	(0.0058)
Lean	-0.0936*	-0.0628	0.7248	0.3364	-0.0034	-0.0166	-0.0121	-0.0200
	(0.0561)	(0.0487)	(0.8861)	(0.7853)	(0.0096)	(0.0107)	(0.0148)	(0.0180)
Posting group 1 (names)	-0.0499	-0.0072	-0.0830	-0.4096	-0.0081	-0.0076	0.0038	0.0007
	(0.0553)	(0.0325)	(0.9339)	(0.7730)	(0.0101)	(0.0095)	(0.0131)	(0.0124)
Posting group 2 (IDs)	-0.1021*	-0.0355	1.4287*	0.2472	-0.0011	0.0031	-0.0164	-0.0121
	(0.0614)	(0.0438)	(0.8018)	(0.7815)	(0.0104)	(0.0109)	(0.0142)	(0.0170)
Constant	0.9776***	1.7193***	90.3028***	87.4869***	0.0953***	0.2802***	0.2588***	-0.1257***
	(0.0428)	(0.1134)	(0.5702)	(1.3161)	(0.0071)	(0.0289)	(0.0085)	(0.0440)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	310084	310084	310084	310084	310084	310084	310084	310084
Adjusted R-squared	0.008	0.128	0.010	0.052	0.006	0.048	0.004	0.107

Table A8: Date and driver fixed effects - long window

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.0829*	0.0822*	-1.4671	-1.4290	0.0264*	0.0228	0.0375**	0.0322**
	(0.0423)	(0.0426)	(1.0750)	(1.0707)	(0.0154)	(0.0148)	(0.0154)	(0.0144)
Post*Posting group 2*Lean	0.0350	0.0333	-1.6876*	-1.6902*	-0.0026	-0.0013	0.0127	0.0154
	(0.0399)	(0.0402)	(1.0178)	(1.0144)	(0.0152)	(0.0148)	(0.0147)	(0.0139)
Post*Posting group 1	-0.0516**	-0.0534**	0.7238	0.7138	-0.0095	-0.0074	-0.0251**	-0.0209**
	(0.0259)	(0.0254)	(0.6759)	(0.6738)	(0.0070)	(0.0069)	(0.0097)	(0.0097)
Post*Posting group 2	0.0095	0.0111	-0.0248	-0.0191	0.0123	0.0107	0.0037	0.0005
	(0.0281)	(0.0276)	(0.7814)	(0.7782)	(0.0108)	(0.0105)	(0.0098)	(0.0096)
Post*Lean	-0.0310	-0.0319	0.4345	0.4371	0.0020	0.0023	-0.0128	-0.0119
	(0.0283)	(0.0288)	(0.7643)	(0.7586)	(0.0097)	(0.0093)	(0.0104)	(0.0096)
Posting group 1*Lean								
Posting group 2*Lean								
Post	-0.0031	-0.0047	1.0660**	1.0748**	0.0069	0.0069	-0.0033	-0.0025
	(0.0194)	(0.0187)	(0.5220)	(0.5212)	(0.0074)	(0.0074)	(0.0065)	(0.0065)
Lean								
Posting group 1 (names)								
Posting group 2 (IDs)								
Constant	0.9637***	1.8205***	91.4879***	86.2097***	0.1193***	0.1670***	0.2628***	-0.0866***
	(0.0126)	(0.0411)	(0.3789)	(0.6108)	(0.0057)	(0.0105)	(0.0054)	(0.0184)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	310084	310084	310084	310084	310084	310084	310084	310084
Adjusted R-squared	0.515	0.527	0.526	0.529	0.264	0.277	0.430	0.483

Table A9: Matched analysis with date and driver fixed effects - long window

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.0990*	0.1004*	-0.5058	-0.4726	0.0445**	0.0389**	0.0464**	0.0378*
	(0.0549)	(0.0541)	(1.4009)	(1.3983)	(0.0177)	(0.0175)	(0.0199)	(0.0194)
Post*Posting group 2*Lean	0.0031	0.0099	-1.1112	-1.1595	-0.0131	-0.0117	0.0040	0.0025
	(0.0466)	(0.0460)	(1.4296)	(1.4241)	(0.0156)	(0.0156)	(0.0165)	(0.0166)
Post*Posting group 1	-0.0566	-0.0582	-0.2796	-0.2884	-0.0203**	-0.0179*	-0.0290**	-0.0249*
	(0.0361)	(0.0354)	(0.7970)	(0.7936)	(0.0093)	(0.0091)	(0.0137)	(0.0135)
Post*Posting group 2	0.0319	0.0299	-0.2026	-0.1741	0.0198	0.0174	0.0095	0.0071
	(0.0359)	(0.0353)	(1.0579)	(1.0549)	(0.0134)	(0.0134)	(0.0124)	(0.0123)
Post*Lean	-0.0375	-0.0439	0.0202	0.0464	-0.0054	-0.0041	-0.0176	-0.0126
	(0.0354)	(0.0341)	(1.1392)	(1.1336)	(0.0082)	(0.0084)	(0.0128)	(0.0130)
Posting group 1*Lean								
Posting group 2*Lean								
Post	0.0041	0.0048	1.5634**	1.5626**	0.0112	0.0108	0.0007	-0.0002
	(0.0245)	(0.0231)	(0.6650)	(0.6673)	(0.0097)	(0.0096)	(0.0086)	(0.0086)
Lean								
Posting group 1 (names)								
Posting group 2 (IDs)								
Constant	0.9436***	1.7822***	91.8457***	86.7289***	0.1204***	0.1765***	0.2631***	-0.0862***
	(0.0167)	(0.0558)	(0.5084)	(0.7706)	(0.0074)	(0.0129)	(0.0069)	(0.0233)
Controls	N	Y	N	Y	N	Y	N	Y
Date and Driver FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	198831	198831	198831	198831	198831	198831	198831	198831
Adjusted R-squared	0.519	0.531	0.527	0.529	0.261	0.274	0.437	0.488

Table A10: Matched analysis - long window

Dependent variable:	Log(Gap Score)		Shift Score		Log(Idle Time)		Log(Fuel Lost)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Posting group 1*Lean	0.1230**	0.1366**	-0.6115	-0.5885	0.0573***	0.0536***	0.0645***	0.0528***
	(0.0570)	(0.0522)	(1.4459)	(1.3966)	(0.0186)	(0.0176)	(0.0196)	(0.0187)
Post*Posting group 2*Lean	0.0116	0.0290	-0.9394	-0.9843	-0.0025	0.0012	0.0125	0.0097
	(0.0582)	(0.0550)	(1.4392)	(1.3895)	(0.0188)	(0.0172)	(0.0227)	(0.0199)
Post*Posting group 1	-0.0563	-0.0702**	-0.2320	-0.1899	-0.0238*	-0.0228*	-0.0333***	-0.0281**
	(0.0365)	(0.0324)	(0.7581)	(0.7241)	(0.0129)	(0.0116)	(0.0123)	(0.0118)
Post*Posting group 2	0.0429	0.0304	-0.1244	-0.1089	0.0216	0.0176	0.0063	0.0070
	(0.0452)	(0.0429)	(1.0250)	(0.9792)	(0.0168)	(0.0156)	(0.0178)	(0.0150)
Post*Lean	-0.0524	-0.0629*	-0.2213	-0.1467	-0.0148	-0.0152	-0.0250	-0.0197
	(0.0412)	(0.0370)	(1.1712)	(1.1342)	(0.0105)	(0.0091)	(0.0151)	(0.0141)
Posting group 1*Lean	0.0791	0.0207	0.0990	-0.6190	0.0041	0.0015	-0.0115	0.0110
	(0.1113)	(0.0900)	(1.5884)	(1.4769)	(0.0229)	(0.0221)	(0.0247)	(0.0317)
Posting group 2*Lean	-0.0633	0.0739	0.9075	-0.0142	-0.0068	0.0076	-0.0311	0.0371
	(0.1058)	(0.0798)	(1.5747)	(1.3542)	(0.0176)	(0.0160)	(0.0265)	(0.0301)
Post	0.0626**	0.0735***	1.7184***	1.6373***	0.0265***	0.0261***	0.0308***	0.0266***
	(0.0277)	(0.0229)	(0.6430)	(0.6093)	(0.0087)	(0.0079)	(0.0093)	(0.0083)
Lean	0.0007	-0.0004	-0.3159	0.3830	-0.0024	-0.0073	0.0173	-0.0074
	(0.0786)	(0.0651)	(1.2546)	(1.1134)	(0.0124)	(0.0114)	(0.0168)	(0.0246)
Posting group 1 (names)	-0.0248	0.0284	0.7930	0.7404	-0.0028	-0.0092	0.0014	0.0145
	(0.0701)	(0.0604)	(1.0287)	(0.9204)	(0.0149)	(0.0165)	(0.0153)	(0.0231)
Posting group 2 (IDs)	0.0266	-0.0850	0.7728	1.0107	0.0050	-0.0025	0.0024	-0.0339
	(0.0761)	(0.0630)	(1.0768)	(1.1107)	(0.0126)	(0.0126)	(0.0181)	(0.0240)
Constant	0.8986***	1.5815***	91.0687***	88.5450***	0.0929***	0.2807***	0.2459***	-0.1567***
	(0.0562)	(0.1392)	(0.7216)	(1.5272)	(0.0086)	(0.0313)	(0.0106)	(0.0503)
Controls	N	Y	N	Y	N	Y	N	Y
Observations	198831	198831	198831	198831	198831	198831	198831	198831
Adjusted R-squared	0.008	0.130	0.006	0.051	0.008	0.051	0.005	0.108