PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Foreign, Commonwealth & Development Office (FCDO) funded programme run by the Centre for Economic Policy Research (CEPR).

Monitoring in Target Contracts: Theory and Experiment in Kenyan Public Transit

Erin M. Kelley Gregory Lane David Schönholzer*

October 23, 2020

Abstract

We develop a relational contracting model to study the role of monitoring in firms and evaluate the model experimentally in the field. Specifically, we introduce monitoring devices into commuter minibuses in Nairobi, Kenya, that track real-time vehicle driving behavior and daily productivity. We randomize which minibus owners have access to these monitoring data using a novel mobile app that we designed for the industry. In line with model predictions, we find that treated vehicle owners modify the terms of the contract by decreasing the transfer they demand in exchange for higher effort and lower risk-taking. Drivers respond accordingly by working more hours and decreasing risky driving behavior associated with higher repair costs. As a result, firm costs fall and profits increase. Structural estimation via simulated method of moments demonstrates a close match of the data to the contract model and suggests overall welfare increases stemming from lower firm costs.

*Kelley: Development Impact Evaluation Unit, World Bank (erinmkelley@worldbank.org). Lane: Department of Economics, American University (glane@american.edu). Schönholzer: Institute for International Economic Studies, Stockholm University (david.schonholzer@iies.su.se). We are grateful to Fiona Burlig, Alain de Janvry, Kyle Emerick, Fred Finan, Nancy Gallini, Jeremy Gordon, John Loeser, Jeremy Magruder, Dave McLaughlin, Edward Miguel, Matt Pecenco, Marco Gonzalez-Navarro, Kate Pennington, Louis Preonas, Manaswini Rao, Elisabeth Sadoulet, Deirdre Sutula, Ralph Winter, and Matt Woerman for helpful comments and suggestions. We gratefully acknowledge financial support from DIL and PEDL. We thank Jeremy Gordon for designing an amazing tracking device. We are also extremely grateful to the entire team at EchoMobile (Boris, James, Vitalis, Langat, Wambururu, and Kevin) for exceptional field management and implementation as well as Peter Waiganjo Wagacha for guidance and support. All errors are our own. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. This paper was previously circulated under the name "The Impact of Monitoring Technologies on Contracts and Employee Behavior: Experimental Evidence from Kenya's Transit Industry".

1 Introduction

A fundamental insight of contract theory is that asymmetric information between firms and employees constrains firm productivity. If firms cannot observe relevant dimensions of worker behavior, contracts fall short of inducing profit-maximizing behavior (Holmström, 1979; Grossman and Hart, 1983; Hart and Holmström, 1986). The recent rise in information technologies may alleviate moral hazard by enabling firms to monitor employee behavior (Hubbard, 2000, 2003). But how do these technologies affect contracts and worker behavior in practice, particularly in environments with weak contract enforcement? And do these effects translate into higher firm productivity?

These questions are central to our understanding of how moral hazard and monitoring affect firm productivity, especially in low income countries where low firm productivity is pervasive (Bloom et al., 2010). But they are difficult to answer. The adoption of monitoring technologies is nonrandom, rendering comparisons across firms uninformative. Even with an appropriate identification strategy, we often lack fine-grained data to observe any changes that may take place within firms once monitoring is introduced. Finally, interpreting the impact of monitoring is challenging because it can trigger subtle strategic responses by both firms and employees.

We overcome these challenges in the context of Kenya's informal public transit industry. This industry is similar to public transport operations in many low and middle income countries, representing millions of travelers and commuters worldwide. The dynamics in informal transit are also prevalent in other industries where output is hard to observe and workers are poor. In Kenya, informal transit is dominated by privately owned minibuses called matatus. The matatu owner hires a driver on an informal daily contract, setting a revenue target for the driver to transfer at the end of the day; the driver retains the residual revenue; and the owner is liable for major expenses accrued during the day – a classical principal-agent setting.

To understand how monitoring affects this contractual arrangement, we develop a new monitoring system tailored to the industry that tracks the driver's effort and risk-taking choices. Specifically, the system allows to observe the driver's location, hours worked, distance driven, and a number of safety violations. We fit 255 matatus with these tracking devices, working exclusively with owners who only manage a single matatu.[1] We then conduct a randomized control trial (RCT) where we provide half of the owners with access to the monitoring system for six months, while the other half continues to manage their drivers according to the status quo.

We find that treated owners are able to effectively use the system to monitor their driver's activities more easily. While they retain the target contract structure, they slightly lower the driver's daily revenue target. Treated drivers work longer hours and engage in substantially less

[1] About one third of owners in our baseline sample of the industry in Nairobi owned only a single matatu.

damaging driving behavior such as off-road driving. This lowers maintenance costs and contributes to substantial increases in daily profit for the owners. After six months, we observe treated owners expanding and investing in their minibus business.

To interpret the findings from this experiment, we develop a principal-agent model that accounts for the contracting constraints faced in such settings. First, as in many principal-agent models, the owner cannot observe effort and risk. Second, in contrast to much of the literature, the owner cannot observe the amount of revenue the driver collects, and can only rely on the transfer from the driver to determine whether to rehire him for the next day. Third, drivers are often liquidity constrained and thus the contract is subject to limited liability. Finally, the lack of contract enforcement requires contracts to be self-enforcing. We study the optimal contract in this environment and the effects of monitoring on employee and firm outcomes.

Our model delivers several important insights into the nature of contracting in these environments. Because revenue is unobserved and limited liability constraints binds, we find that the optimal contract hinges on a single parameter: the daily revenue target, which determines a simple linear rehiring rule based on the amount transferred. This contract incentivizes excessive risk-taking *and* high effort, and it is inefficient relative to a social planner benchmark. The principal cannot contract lower risk-taking in a way that is both incentive compatible for the driver and maintains the flow of transfers from the driver to the owner. Monitoring technologies make effort and risk observable to the owner, allowing the owner to specify the amount of effort and risk they want the driver to supply. As a result, profits rise primarily from less risk-taking which results in fewer costs to the firm. Welfare effects of monitoring are ambiguous, and rather surprisingly, it is possible for the driver's utility to rise.

To quantify owner and driver welfare under the status quo and with the introduction of monitoring, we estimate the structural parameters of the contract model using data from our experiment. We first estimate the disutility of driving choices, firing costs, the value of the driver's outside option, and owner welfare via simulated method of moments (SMM) with data from the control group. The model fits the data well: estimates closely match the observed values for the daily target, the expected profit, the driver's residual revenue, the probability of separations, and the driver's valuation of his outside option. Our estimates suggest the present-discounted contract surplus is large: the driver values the contract at about $442 and the owner at about $1,116.

We then apply our SMM procedure to estimate the welfare effects of monitoring. Matching on the reduced-form effect of monitoring on the target, the expected profit, the expected residual revenue, and the probability of separation, we estimate that the driver's present-discounted value of the contract falls by $31 or about 7% due to monitoring. This is primarily because they incur 20% more disutility from the effort-risk bundle under monitoring, although drivers report improvements

in their relationship with owners. On the other hand, the owner gains about $41 from higher profits under similar revenue. This is very similar to the average willingness to pay of $45 for the device at the end of the study. Thus, the total welfare effect of monitoring is positive but small.

Our study contributes to a number of literatures. The paper speaks to the vast theoretical literature on principal-agent relationships and contract formation in firms, which predominantly focuses on deriving the optimal contracts subject to various constraints (Holmström, 1979; Grossman and Hart, 1983; Hart and Holmström, 1986; Innes, 1990; Levin, 2003). Removing information asymmetries is often an immediate path to improving productivity in these models. Our paper provides a theoretical and experimental evaluation of how contracts change when these constraints are alleviated, but not entirely removed, via monitoring technologies.

We build on seminal empirical work by Hubbard and (2000; 2003) and Baker and Hubbard (2004) who investigate how the introduction of onboard diagnostic computers affected the U.S trucking industry. They find that monitoring increases productivity by raising the returns to delegation, increases efficiency, and leads to firm growth. In contrast to these studies, we generate exogenous variation in the usage of monitoring technologies by randomizing which companies receive data from a tracking device. We also capture high frequency data on contracts and worker behavior. This allows us to monitor how different dimensions of the contract and worker behavior change over time and how these changes affect firm profit.

There is also some evidence to suggest that monitoring systems can have differential impacts on firm dynamics in low-income countries. The quality of management practices is typically lower (Bloom et al., 2013), which could prevent firms from harnessing the benefits of monitoring technologies. Moreover, contract enforcement is weak and workers are poor, which could also limit the impact of these systems. Our work demonstrates that this is not the case, but underscores how the impact of monitoring differs systematically in low-income countries because of these constraints. In particular, we do not find changes in the ownership structure (unlike Baker and Hubbard 2004) – contract change is more subtle, reinforcing Macchiavello and Morjaria (2015) findings about relational contracts. In closely related contemporary work, de Rochambeau (2020) finds that monitoring induces Liberian truck drivers to supply higher effort, although her focus is on intrinsic motivation and driver heterogeneity rather than the contract. Other work on monitoring in developing countries studies primarily external monitoring of firms and public service providers (Duflo et al., 2012; Björkman and Svensson, 2009; Duflo et al., 2013).

Our work also speaks to a large literature documenting barriers to firm growth in developing countries. Empirical research on small firm growth has identified three key challenges firms face: credit constraints, labor-market frictions, and managerial deficits (Bloom et al., 2010). Our paper most closely resembles the work on managerial deficits, which refers to the difficulties firms face

managing day-to-day operations (including financial accounts and inventories), and incentivizing and monitoring workers. Most of the work in this field studies the impact of business training programs (Bloom et al., 2013; Berge et al., 2015; McKenzie and Woodruff, 2017). These interventions train firms about how to manage aspects of the business that do not involve employees, such as maintaining business records, separating finances, inventory, controlling for quality, or marketing. In contrast, our paper focuses on employee management through informal contracts, the role of moral hazard, and how providing information about employee behavior can change firm operations.

Finally, there is growing recognition in economics that the informal transit industry has a large impact on various development outcomes. In many low and middle income countries, informal transit systems such as matatus provide essential infrastructure in the absence of formal, state-run mass transit, but they are often inefficient and unsafe (Cervero and Golub, 2007). Habyarimana and Jack (2011, 2015) study the role of leveraging passenger monitoring of driving behavior through evocative messages in matatus. Our work provides the first systematic analysis of the most prevalent contractual arrangement in this industry and how monitoring technologies may help the industry become more efficient as these devices become more widespread.

The rest of the paper is organized as follows. We describe features of the informal transit industry generally and specifically for Nairobi in Section 2. Section 3 then develops a theory of contracting in this industry. In Section 4 we describe the monitoring technology, the data collection, and the experimental design. We present reduced-form results of the experiment in Section 5. Section 6 provides results from our structural estimation, and Section 7 concludes.

2 Context: The Matatus of Nairobi

Informal transit systems are the backbone of public transportation in low-income countries, often comprising more than two thirds of commutes (Godard, 2006).[2] Private entrepreneurs typically own a small fleet of vehicles and operate on defined routes that are managed by cooperatives. Passengers embark at various points along the route and pay in cash (Bruun and Behrens, 2014). Drivers are commonly hired on target contracts: the owner sets a daily revenue target at the beginning of the day and the driver is the residual claimant.[3] If the driver misses the target, the owner typically expects to receive the full day's revenue. If the owner deems the transfer to be too low, she will reconsider whether to rehire the driver for the next day. The owner sets the target based on vehicle characteristics, the route, and day-specific shocks such as weather conditions or special events (e.g

[2]Informal transit refers to all forms of privately-run public transportation services such as minibuses, vans, taxis, station wagons, three-wheelers, or motorcycles (Cervero and Golub, 2007).

[3]In our context, the driver is accompanied by a fare collector who works with the driver as a team. For the purposes of this study, we treat them as a unit. The norm in Kenya is for them to split the residual revenue evenly.

the beginning of the school year) (Behrens et al., 2015).

In Kenya, informal transit services are essential. Rough estimates suggest that 15,000 private minibuses circulate throughout the city. The industry employs over 500,000 people and contributes up to 5% of the country's GDP (Kenya Roads Board, 2007). Private entrepreneurs primarily purchase 14 seat minibuses, and license their vehicles to a particular route managed by a cooperative. Owners hire a single driver to operate their vehicle along the designated route. An owner's day-to-day management consists of calling their driver, checking whether the vehicle needs to be serviced, and occasionally staging observers along the route to learn about the driver's activities.

While this informal network of buses constitutes the only dependable transit system in Nairobi, the industry is widely perceived to be suffering from a number of inefficiencies (McCormick et al., 2013; Behrens et al., 2015; Mutongi, 2017). For example, the presence of severe competition within a route leads to reckless driving and high vehicle maintenance costs. Similarly, a lack of regulation and enforcement creates incentives for drivers to operate on unlicensed routes, where they pay substantial fines when they are caught. Owners are unable to limit these events because they cannot easily observe their driver's activities – a management problem that deters owners from expanding their fleet and growing their business.

In recent years, companies offering GPS tracking services have started to enter the Kenyan market. Many insurance providers have also mandated that minibuses install GPS trackers for security reasons.[4] Despite their increasing availability, most medium and small-scale minibus owners have not installed them in their own vehicles because they are prohibitively expensive (around $600 per unit). To fill this need, we worked with a Kenyan technology company to create a new monitoring system that is considerably cheaper and more flexible than other tracking systems. We describe the system in detail in Section 4.

Kenya's transportation industry is appealing to study for a number of reasons. First, it is broadly representative of informal transit systems around the world. Second, the dynamics we study between firms and employees can also be found in other industries where revenue is collected by the employee and remains unobserved by the firm. Finally, this context allows us to overcome major data constraints that have limited researchers ability to study how contracts respond to monitoring technologies in the real world.

[4]See Business Daily Africa article "Vehicle tracking system gains popularity in Kenya", September 22, 2009, `https://www.businessdailyafrica.com/corporate/539550-661514-6k6toiz/index.html`.

3 Model

We now describe a contract model of the owner-driver relationship in the informal transit industry. The defining features of this model are not unique to our setting and reveal important insights about a class of contracts that are prevalent in low-income countries. The goal of this model is threefold. First, it allows us to precisely describe the mechanics that lead the contract to be inefficient. Second, we can derive strong predictions about the effect of monitoring on the principal, the agent, and firm outcomes. Finally, the model provides the basis for the structural estimation of welfare under the baseline contractual arrangement and after monitoring is introduced.

To accurately reflect the informal transit environment, we combine several model components from the contract theory literature. Since drivers are relatively poor, we include a limited liability constraint as in Innes (1990). Because contract enforcement is limited, we require the driver's commitment to the contract to be self-enforcing, as in Levin (2003). The most novel component is to make output (or revenue) unobservable to the owner in addition to driver's choices. While this echoes the idea of costly state verification introduced by Townsend (1979), it generates new and interesting dynamics pertinent to the informal transit industry as well as other environments where the principal struggles to observe output.

We begin by setting up the model in the baseline environment without monitoring and derive an intermediate result that greatly simplifies the model's solution. We then show how the resulting contract compares to a social-planner benchmark in the form of an integrated owner-driver for whom the agency problem is of no concern. Finally, we show how the contract changes when monitoring technologies are introduced, which allows the owners to observe some driver choices. We refer to the principal as the female owner (of the vehicle) and the agent as the male driver throughout the model.

3.1 Setup

Technology and Preferences A risk-neutral owner (principal) and risk-neutral driver (agent) engage in a daily relational contract. They value the contract at endogenous values V and U, respectively, and discount the future with a common factor δ. The driver chooses effort e and risk r during the day from $\mathcal{S} = [0, \bar{e}] \times [0, \bar{r}]$ where $\bar{e} > 0$ and $\bar{r} > 0$, incurring disutility $\psi(e, r) \in [0, \bar{\psi}]$. On the basis of these actions, nature draws revenue $y \in \mathcal{Y} = [0, \bar{y}]$ from the revenue distribution $G(\cdot|e, r)$ as well as repair costs $c(r) \in [0, \bar{c}]$ from $F(\cdot|r)$. Repair costs depend on risk but not effort and accrue entirely to the owner.[5] Conditional on effort and risk, the revenue and cost distributions

[5]Note that the role of risk is nonstandard in the literature, compared to, for example Ghatak and Pandey (2000). In our model, risk is essentially a second effort dimension which has an additional cost to the principal, instead of having a mean-preserving effect on the variance of output. We nonetheless call this choice "risk" because it

are independent. We make the following assumptions about the functional forms of technology and preferences:

Assumption 1. *Technology and preferences obey:*

- *$G(y|e,r)$ is twice continuously differentiable with respect to all arguments with density $g(y|e,r)$.*
- *The distribution of revenue y has the monotone likelihood ratio property (MLRP), which implies the first-order stochastic dominance (FOSD) property: for all y, we have that $G_s(y|e,r) \leq 0$ for $s \in \{e,r\}$.*
- *We assume that effort and risk are weak substitutes with $G_{er}(y|e,r) \geq 0$ and $G_{ss}(y|e,r) > 0$ for $s \in \{e,r\}$.*
- *Driver disutility is twice continuously differentiable with partials $\psi_s(e,r) > 0$ for all $(e,r) > 0$, $\psi_{er}(e,r) > 0$ and $\psi_{ss}(e,r) > 0$ for $s \in \{e,r\}$.*

Information and Contract At the end of the day, the driver transfers $t(y)$ and keeps the residual $y - t(y)$. She then rehires the driver with some probability $p(\cdot)$. In the baseline environment without monitoring, this rehiring schedule depends only on the transfer: $p(t)$. If the owner has access to the monitoring technology, she can directly observe the driver's effort and risk choices and may use these in the rehiring schedule $p(t,e,r)$. In contrast to standard contracting problems, the owner does not receive any information about revenue, even with a monitoring device.[6] In case the driver gets fired he receives his outside option while the owner pays a hiring cost h before drawing an identical driver.[7]

To summarize the structure of the model, the timing of the game is as follows. At the beginning of the day, the owner and driver agree on contract terms. The driver then makes driving choices $(e,r) \in \mathcal{S}$ during the day. Based on (e,r), nature draws revenue y as well as repair cost c. The driver then transfers $t(y)$ to the owner and keeps $y - t(y)$. If the driver is indifferent between two transfers, we assume he will transfer the higher amount. Finally, the owner rehires the driver for the next day with probability $p(t)$, or $p(t,e,r)$ in case of monitoring.

corresponds closely to actions such as a risky maneuver to overtake another car in traffic; driving offroad to bypass traffic, risking damage to the vehicle; or taking an unlicensed route, risking a fine.

[6] This is an important feature of monitoring in informal transit. Even if the owner knows the exact number of trips a driver took there is no way to get a precise estimate of revenue because they do not know the number of passengers. Drivers said they were more comfortable with GPS technologies that revealed their choices of effort and risk, as opposed to technologies that allow owners to observe revenue directly, such as electronic payment systems (such as BebaPay, a failed Google venture).

[7] For expositional simplicity, we set $h = V$ in the theoretical section, but we estimate the hiring cost separately in the structural estimation. Our focus in this paper is moral hazard in stable owner-driver relationships so we do not study the interesting but separate problem of adverse driver selection.

Baseline contracting problem without monitoring In the status quo contracting problem, the owner maximizes the expected sum of transfers and the future discounted value of the contract minus the cost of risk subject to constraints:

$$V = \max_{e,r,t(y),p(t)} \mathbb{E}\left[t(y) - c(r) + p(t(y))\delta V | e, r\right] \tag{1}$$

subject to

1. $U = \mathbb{E}\left[y - t(y) + p(t(y))\delta U | e, r\right] - \psi\left(e, r\right) \geq \bar{u}$
2. $(e, r) \in \arg\max_{(\tilde{e},\tilde{r})\in\mathcal{S}} \mathbb{E}\left[y - t(y) + p(t(y))\delta U | \tilde{e}, \tilde{r}\right] - \psi\left(\tilde{e}, \tilde{r}\right)$
3. $t(y) \leq y$
4. $y - t(y) + p(t(y))\delta U \geq y$
5. $t(y) \in \arg\max_{\tilde{t}\geq 0} y - \tilde{t} + p(\tilde{t})\delta U$,

where the owner's expectation is over the joint distribution of y and c. While the driver ultimately chooses effort and risk, the owner treats them as choice variables for the purpose of designing the contract, as is standard in contract theory. The first two constraints are standard participation constraints and (driving) incentive compatibility constraints. Driver utility is the expected sum of the residual revenue and the future discounted value of the contract minus the disutility of effort and risk. The participation constraint restricts driver utility to be at least as great as his outside option. For now, we assume that this outside option is normalized to zero, $\bar{u} = 0$, which we relax in our structural estimation. The third constraint is the limited liability constraint, which restricts the driver from transferring more to the owner than what he made on a given day. The fourth constraint ensures dynamic enforceability: the driver has to prefer to honor the terms of the contract ex post over reneging. The fifth and last constraint restricts the transfer to the owner to be incentive compatible: $t(y)$ has to be an optimal transfer from the driver's point of view.[8]

Although reminiscent of a fixed rental contract, the resulting target contract is structurally different from known contracts in the literature. Limited liability prevents the driver from paying a rental price upfront. Hence the owner has to rely on a transfer at the end of the day based on

[8]Since firing the driver is costly to the owner, she may have an incentive to renege on the agreed-upon rehiring probability $p(t)$ and rehire him despite a negative outcome of the rehiring lottery. For simplicity, we do not explicitly model this possibility. It would require the driver to form beliefs about the likelihood that the owner will renege, and then for the owner to take this into account when considering the contract. While it may be possible to incorporate this incentive into the model, we are likely to arrive at similar conclusions in terms of contract dynamics with respect to driver choices and the transfer problem. For the contract not to unravel, we assume that frequent reneging would be inferred over time by the driver and he would switch to a strategy of transferring nothing to the owner.

uncertain revenue. This contract structure is not unique to our setting. It is common whenever workers are relatively poor and output is hard to observe.

3.2 Linear Rehiring and Transfer Schedules

The solution to this contracting problem can be greatly simplified with an assumption about the driver's risk preferences relative to the owner. To this end, we define the *incentive-compatible set*

$$\mathcal{I} = \left\{ (e,r) \in \mathcal{S} : \frac{\int_0^{\bar{y}} y g_e\left(y|e,r\right) dy}{\int_0^{\bar{y}} y g_r\left(y|e,r\right) dy} = \frac{\psi_e\left(e,r\right)}{\psi_r\left(e,r\right)} \right\}$$

as the optimal effort-risk bundles that the driver would choose for a given disutility budget – he equates the ratio of marginal benefits from an additional unit of effort and risk to the ratio of marginal costs. See Panel A of Figure 2 for an illustration. Among these bundles, we can then define the driver's bliss point to be:

$$(e_D, r_D) = \arg\max_{(e,r)\in\mathcal{S}} \mathbb{E}\left[y|e,r\right] - \psi(e,r),$$

which is the driver's preferred effort-risk bundle if he were to find the vehicle at the side of the road and did not have to worry about repair cost or contract concerns. We then make the following assumption:

Assumption 2. *Relative costliness of risk: At any bundle* $(e,r) \in [e_D, \bar{e}] \times [r_D, \bar{r}]$*, the owner prefers less risk than the driver.*

This assumption reflects the prevalent sense among owners that drivers engage in excessive risk-taking. The owner would prefer an effort-risk bundle skewed more towards effort because the costs of risk-taking accrues exclusively to them. See Panel B of Figure 2 for an illustration of this assumption: the owner's indifference curves in blue reflect the fact that she always prefers higher effort but faces a tradeoff with respect to risk. Higher risk-taking increases revenue and the expected transfer she receives, but it also increases repair costs. The assumption states that the owner prefers the driver to choose a bundle with less risk because at the driver's bliss point the costs of risk outweigh the potential benefits.

Under this assumption, the rehiring schedule collapses to a single parameter which acts as a contractual shorthand in our setting: the daily target T, above which the next day's re-employment is guaranteed. In contrast, without this assumption, the rehiring schedule would depend on the functional form of revenue. The target is also a key parameter in our structural estimation below. Specifically, we can state:

Lemma (Minimal linear contract)**.** *Under Assumptions 1 and 2, in any solution to the baseline contracting problem without monitoring, the following schedules are optimal:*

$$t(y) = \min\{y, T\}$$

and

$$p(t) = p_0 + \frac{t}{\delta U}$$

for $t \leq T = (1 - p_0)\,\delta U$ *and* $p(t) = 1$ *for* $t > T$*, where* $p_0 \in [0, 1]$*.*

We provide a proof of the Lemma in Appendix A.1. Under an optimal contract, the transfer schedule $t(y)$ requires the driver to transfer all revenue up to some target amount T, defined as the transfer at which re-employment is guaranteed. The driver retains any revenue beyond T. The corresponding rehiring schedule is linear up to the target, where it reaches certainty.

The intuition for these schedules follows from the various goals the owner pursues. First, she seeks to maximize the transfer for any given revenue the driver achieves. To this end, the rehiring schedule must guarantee that the marginal benefit to the driver of one additional dollar transferred (which is just the change in the rehiring probability times the discounted value of that relationship $p'(t)\delta U$) exceeds the direct value of keeping that dollar (which is just 1). This means the slope of the rehiring schedule needs to surpass the inverse of the discounted value of the relationship $(1/\delta U)$, as illustrated in Figure A.3.

Second, the owner seeks to incentivize the driver to select her preferred level of effort and risk. Since she cannot observe driving choices, she can only induce effort-risk bundles on the incentive compatible set (see Panel A in Figure 2). This means her choice comes down to bundles with both higher effort and higher risk, or bundles with lower effort and lower risk. If she sets the slope of the rehiring schedule to $1/\delta U$, the driver chooses his bliss point. She *could* induce higher effort-risk bundles by setting a rehiring schedule that is steeper than $1/\delta U$. The driver would then find effort and risk more appealing because of its high return in terms of increased rehiring probability in case he misses the target. However, she *cannot* induce effort-risk bundles below the driver's bliss point by setting the rehiring slope below $1/\delta U$ because the driver would keep the marginal dollar rather than transfer it, without actually lowering effort and risk. Since the owner prefers less risk than the driver according to Assumption 2 (Panel B of Figure 2), she contents herself with the lower bound of risk induced by the minimal slope $1/\delta U$ and resigns herself to capturing as much revenue as possible. As shown in Panel C of Figure 2, this is her preferred bundle among those that are both incentive compatible and transfer compatible.

Finally, the owner also needs to satisfy dynamic enforcement and limited liability, both of which are automatically satisfied under the linear rehiring and transfer schedules.

3.3 Baseline Optimal Contract

Armed with the simplifications provided by the Lemma, we can now assess the efficiency of the baseline contract relative to a natural benchmark: the optimal decision of an integrated decision maker, taking into account both the repair costs due to risk as well as the disutility of effort and risk. We refer to this benchmark as the social planner or owner-driver solution, since this corresponds to the decision problem of an owner who drives her own vehicle.

Specifically, the social planner ("owner-driver") problem is

$$(e^*, r^*) \in \arg\max_{(e,r)\in\mathcal{S}} \mathbb{E}\left[y - c(r) + \delta W | e, r\right] - \psi\left(e, r\right)$$

where W is the owner-driver's continuation value. The baseline contract without monitoring compares as follows to the social planner's solution:

Proposition 1 (Inefficiency of baseline contract)**.** *Let (e_B, r_B) be the driver's baseline effort-risk profile without monitoring. Under Assumptions 1 and 2, the following properties hold in the baseline contract:*

- *The driver takes excessive risk $r_B > r^*$.*
- *Effort may be over- or undersupplied: $e_B \lesseqgtr e^*$.*
- *Welfare is suboptimal: $\mathbb{E}\left[y - c(r_B)|e_B, r_B\right] - \psi\left(e_B, r_B\right) < \mathbb{E}\left[y - c(r^*)|e^*, r^*\right] - \psi\left(e^*, r^*\right)$.*

See Appendix A.2 for a proof. This proposition says that the baseline contract is inefficient compared to the social planner solution due to excessive risk taking by the driver. Risk is oversupplied relative to the social optimum because in choosing r_B the driver is not accounting for repair costs accruing the owner. Unlike many other principal-agent models, effort provision could be too high or too low depending on the degree of substitutability between effort and risk. If effort and risk are weakly substitutable then higher risk could induce higher effort than e^*.

The failure of the contract to achieve the first-best outcome reflects the owner's inability to steer the driver away from his preferred mix of effort and risk. Hence, the owner may be able to use monitoring technologies to overcome this limitation and move the contract closer to the first best. We now turn to examining this possibility.

3.4 Introducing Monitoring

In the *target contract with monitoring* the owner observes the driver's effort and risk choices and conditions her rehiring schedule on them as well as the transfer: $p(t, e, r)$ instead of $p(t)$. In this case, the solution exhibits the following features.

Proposition 2 (Effects of monitoring). *Let (e_M, r_M) be the driver's effort-risk profile under monitoring. Under Assumptions 1 and 2, the target contract with monitoring has the following properties:*

1. *One solution for the rehiring schedule is:*

$$p(t,e,r) = \begin{cases} p_M + \frac{t}{\delta U_M} & \text{if } e = e_M \text{ and } r = r_M \\ 0 & \text{otherwise} \end{cases}$$

2. *Compared to the baseline contract:*

   - *Higher effort provision $e_M > e_B$ and lower risk $r_M < r_B$.*
   - *Revenue may rise or fall: $E[y|e_M, r_M] \lesseqgtr E[y|e_B, r_B]$.*
   - *Profits increase: $\mathbb{E}\left[y - c\left(r_M\right)|e_M, r_M\right] > \mathbb{E}\left[y - c(r_B)|e_B, r_B\right]$.*
   - *The target falls if revenue falls: $T_M < T_B$ if $E[y|e_M, r_M] \leq E[y|e_B, r_B]$.*
   - *The welfare effect is ambiguous:*

$$\mathbb{E}\left[y - c(r_B)|e_B, r_B\right] - \psi\left(e_B, r_B\right) \lesseqgtr \mathbb{E}\left[y - c(r_M)|e_M, r_M\right] - \psi\left(e_M, r_M\right).$$

A proof is in Appendix A.3. Monitoring unambiguously lowers risk. The owner is no longer tied to the incentive compatible set but can now contract an effort-risk bundle that skews the driver towards a more effort-intensive choice. This unambiguously increases profits due to lower expected repair costs. However, expected revenue may be higher or lower: it is possible that the owner settles on an effort-risk bundle that yields lower expected revenue if the corresponding drop in expected repair costs is even larger.

Even with monitoring, the contract retains its target structure. Since monitoring only reveals driver choices but not revenue, the owner must continue to provide transfer incentives, prohibiting the establishment of a wage contract. There are two forces that influence the owner's decision to re-optimize the target. First, she needs to compensate the driver for shifting him away from his incentive compatibility constraint, thereby making him worse off.[9] This puts downward pressure on the target. Second, if expected revenue is higher, the owner increases the target in an effort to capture some of this surplus. However, if expected revenue falls, this reinforces downward pressure

[9] Once the owner has access to monitoring, she internalizes a part of the disutility of effort. This is because the driver's valuation of the contract U falls with higher effort, and according to the Lemma, this implies a lower target, which in turn lowers owner utility. As a result, the owner's utility no longer rises monotonically with effort, as illustrated in Panel D of Figure 2. Note, this also implies that even under monitoring the driver's participation constraint does not bind. This stems from the fact that the owner relies on providing excess value to the driver above their outside option in order to induce transfers.

on the target. Therefore, while the overall impact of the target is ambiguous, we would expect the target to fall if the revenue distribution falls or remains largely the same.

Monitoring may raise or lower overall welfare, depending on whether the contracted effort-risk bundle under monitoring confers higher or lower utility to the driver. While the owner is unambiguously better off, an interesting implication of the contract under monitoring is that the driver can be better off as well. This depends on how much the driver's disutility of driving changes under monitoring: slightly higher disutility under the new effort-risk bundle may be compensated by a lower target, leaving the driver better off. This particular contract was not feasible without monitoring because it was not incentive compatible – the owner could not trust the driver to choose this bundle in exchange for this lower target. With the introduction of monitoring this contract is now enforceable.[10]

4 Experimental Design and Data Collection

To understand the impact of monitoring on the matatu industry, we developed the *SmartMatatu* monitoring system with a Kenyan technology company (Echo Mobile). We developed our own system because available alternatives on the market were either too costly, or not sophisticated enough. The R&D process lasted more than one year, and benefitted from extensive discussions with small-scale matatu owners. The tracking device has a GPS and gyroscope, which capture the vehicle's location and its acceleration at 30 second intervals. The device relies on GPRS to send the information from the tracker to our servers via the cellphone network. The data is further processed on the server to provide daily measures of the vehicles' mileage, the number of hours the vehicle's ignition was on, average and maximal speed, and the number of speeding, over-acceleration, sharp braking and sharp turning alerts.

We also designed a novel mobile application to convey information from the tracker to owners in a user-friendly way (Figure 1). The app's first tab is a map of Nairobi and presents the real-time location of the vehicle. By entering a specific time interval into the phone, the app can display the exact routes traveled by the matatu over this time period. This first tab conveys a more accurate measure of risky driving because owners can see if the driver is operating on bumpy routes. The second tab displays all the safety alerts that are captured by the device. The final tab conveys a summary of the driver's productivity and safety. The productivity section lists the total mileage

[10] To see this more concretely, imagine a point (e, r) on the driver's incentive compatibility set and another point (e_m, r_m) (which we will assume is on the same isoquant for convenience) $e < e_m$ and $r > r_m$. Now imagine that $c(r) >> c(r_m)$, but $\psi(e, r)$ is only slight lower than $\psi(e_m, r_m)$. If the driver could credibly commit to supplying $\psi(e_m, r_m)$, then the owner would optimally choose to lower the target, which would increase driver's utility. However, because T is set before (e, r) are chosen, the owner knows the driver will not follow through on their commitment (which the owner cannot verify), which makes this agreement impossible.

covered, and the duration the ignition was turned on that day. This provides owners with a more accurate measure of driver effort. Finally, the SmartMatatu app was also designed to collect daily information from owners, including the target; the amount the driver transferred; any repair costs incurred; and an overall satisfaction score for their driver's performance.

We conducted an extensive recruitment drive in late 2015 by contacting cooperatives operating across nine major commuter routes in Nairobi. We organized several large meetings with matatu owners, presenting the study's goals and methodology. We registered interested owners that satisfied three conditions: they had to own only a single 14-seater matatu; they had to manage it themselves; and they had to employ a driver rather than driving the minibus themselves. We informed all owners that we would be placing a monitoring device in their vehicle free of charge, and they would be required to provide daily information about their business operations. We also mentioned that a random subset of owners would be selected to receive information immediately, while others would have to wait 6 months before gaining access to the information for a shorter two month period. We successfully registered 255 owners, which we randomized into treatment (126) and control (129).

We conducted installations and trainings from November 2016 to April 2017 (Figure A.4). We installed the trackers under the vehicle's dashboard to prevent tampering. Our field team took owners aside and provided them with an Android smartphone with our SmartMatatu app pre-installed. The app only contained information for owners in the treatment group, who received an additional 30 minutes of training on how to read the information. The field manager then administered a baseline survey and asked the owner to be diligent about submitting the daily survey through the app. The owner received cellphone airtime worth about 40 cents (KES 40) for every submission.

At the same time another enumerator took drivers aside and explained that we were placing a tracking device in the vehicle, and we would be collecting data for research purposes. We did not mention, however, whether the information would be transferred to the owner.[11] We also asked drivers to complete our baseline survey, and we informed them that we would be sending them daily SMS messages to elicit information about the day's operations.[12] Specifically, the message would ask about whether the vehicle was on the road, the amount of revenue they collected, and the residual revenue they kept as a salary. We emphasized that all of the data they shared would remain confidential and they would be compensated 40 cents for each submission.[13]

For the next six months, we sent reminders to owners and drivers to submit data (via the app

[11] This meant that any subsequent changes we observed in driver behavior could only come from owners using the tracker data, rather than from receiving different information from the enumerators during the installation.

[12] If an owner-driver pair separated, we onboarded the driver using the same baseline survey.

[13] We check for differential reporting by looking at the relationship between reported revenue and mileage for treatment and control drivers. We do not find that the relationships are statistically different from one another.

for owners and via SMS for drivers). We offered continued support to treatment owners to help navigate the app. At the end of the six month period, we conducted an endline survey focusing on business investment decisions. We also ran a willingness-to-pay experiment, offering owners two additional months of monitoring information through the app. We then granted control owners access to the information from the tracker for two months. We provide more details on experimental design and data collection in Appendix B.

5 Descriptive Statistics and Experimental Results

We now discuss the reduced-form impacts of the experiment. We fist summarize descriptive statistics about owners and drivers in our sample, and provide evidence for basic features of our contract model. We then turn to a discussion of the impact of the intervention.

5.1 Descriptive Statistics

Baseline Survey and Balance We present baseline sample characteristics and treatment balance in Table 1. Owners are mostly self-employed men in their late thirties. They have eight years of industry experience on average, and five years of experience as matatu owners. Owners have worked with 1.85 drivers on average, and 79% of our sample has worked with at least one driver before the current one. Owners rate their drivers as being fairly honest and diligent. They set a daily revenue target at baseline of approximately $31.30 (KES 3,130) and receive $26 on average from the driver. These characteristics are balanced across treatment and control owners.

Drivers in our sample are exclusively male, they are a few years younger than owners and have lower levels of education. They have eight years of experience as matatu drivers on average, and have worked for six owners in the past. They have spent just over one year with the current owner, and rate the owners as fair. On average, drivers collect approximately $77 (KES 7,700) in passenger fares throughout the day, and keep approximately $9.60 for themselves as a salary. Driver characteristics are balanced, with the exception of driver age and experience (which are strongly correlated). We control for baseline characteristics to account for this imbalance.

The majority of matatu vehicles are imported Japanese minibuses that have been used for thirteen years. Some of them have special features such as free wifi, sound systems, or TVs. They were purchased for approximately $6,675. Appendix Figure A.6 provides a visual of a matatu.

Empirical Contract Characteristics Our data shows that basic elements of the contract align closely with our model. First, we see that the transfer function has the piecewise linear shape we posit in the Lemma: driver transfers increase linearly with revenue until the transfer amount

reaches the target (Figure 3). We interpret the fact that drivers transfer less than total revenue as evidence for an (unmodeled) subsistence income. Note that including this subsistence level would not change the predictions of the model as it would simply constitute a renormalization of the revenue distribution. Therefore, for simplicity we do not include this feature formally in the model. However, we do account for the subsistence income in the structural estimation.[14]

Second, the figure also shows that owner's satisfaction with their driver increases with the size of the transfer, as suggested by the rehiring schedule in the Lemma. While there are only 26 separations between owners and drivers over the course of our study, we find that they occur more frequently when the driver misses the target (Appendix Figure A.7). This provides evidence for the model prediction that owner rehiring is a function of the transfer.

5.2 Experimental Results

We first investigate the degree to which owners engaged with the monitoring app and how the app affected self-reported management practices. We track owners' usage of the device and find that 70% of owners consult the app weekly, while 50% use it daily (Appendix Figure A.5). We see that treated owners find it easier to monitor their drivers than control owners, and spend less time monitoring their drivers (Table 2). We also confirm that owners are internalizing the information we provided through the app. Treated owners are more likely to know about the vehicle's mileage and whether the driver operated on an unlicensed route. They are not more likely to know the vehicle's revenue, which we expect because our monitoring technology does not track the number of passengers who board the vehicle. Finally, treatment owners are also more likely to say that their driver's performance has improved.

To test the predictions of the model summarized in Proposition 2, we run the following regression using daily panel data for vehicle an owner-driver-matatu observation i on day d:

$$y_{id} = \sum_{m=1}^{6} D_{im}\beta_m + \alpha_d + \tau_{r(i)} + \mathbf{X}_i'\gamma + \varepsilon_{id} \tag{2}$$

where y_{id} is an outcome of interest; D_{im} are treatment indicators by month since installation; β_m are our main parameters of interest, the effect of treatment assignment m months after installation; α_d are day fixed effects; $\tau_{r(i)}$ are route fixed effects; $\mathbf{X}_i$ is a vector of baseline characteristics;[15] and

[14] As an alternative to subsistence income, risk aversion on the driver side would lead to a similar pattern in the transfer schedule. However, this would fundamentally complicate the solution the model: the driver's rehiring schedule would become a nonlinear function of the risk aversion parameter. Both the proofs of the propositions and the structural estimation rely on a linear rehiring schedule. Moreover, it is not clear whether introducing risk aversion would have any other benefits besides accommodating the pattern in the transfer schedule.

[15] Specifically, we include as control variables the age of the matatu, the number of special features, owner age and

ε_{id} is an error term, which we allow to be arbitrarily correlated within i across days. This design allows us to examine the treatment effect as it evolves over the six months of the study.[16]

The model predicts that treatment drivers will increase the amount of effort they supply and reduce the amount of risk they take in response to monitoring. This is because the monitoring device allows owners to see the amount of risk and effort drivers choose, and direct drivers towards a more favorable bundle. The results in Figure 4 support this prediction. We proxy driver effort by the number of hours the matatu is operating and find that operating hours increase by 0.98 hours per day on average by the third month after installation and rise steadily until the end of the study.[17] By month six, effort levels increase by 1.47 hours per day on average in the treatment group, a 9.9% increase in drivers' labor supply. This is a substantial increase in an environment where drivers are already working 14-hour days. While this increase in effort leads to gains for the firm, we may worry about safety externalities for passengers and pedestrians exposed to drivers in their 15th hour. We show in a companion paper (Kelley et al., 2020) that there is no significant increase in safety-related outcomes. With more hours on the road, we also see the number of kilometers increase by 12 kilometers per day on average (10%), which corresponds to an extra trip to or from the city center.

Treatment drivers also appear to take substantially less risk. One of the greatest sources of risky driving is when drivers operate on unlicensed routes. Drivers often use these routes as shortcuts to avoid traffic jams where they sit idly without picking up any passengers. These shortcuts are less appealing from the owner's perspective for a number of reasons. The roads are typically less well maintained and bumpier, which means vehicles are more likely to be damaged and repair costs will increase for owners. Furthermore, owners have to pay large fines when drivers are caught along these routes.

Consistent with this reasoning, we find that treated drivers spend less time on these routes after the introduction of monitoring. Panel C in Figure 5 shows that they are about 400 meters closer to the licensed route on average than control drivers throughout the study period. Next, we investigate whether this change in the distance from the licenced route results in less side to side movement. This would indicate that treated drivers are taking less bumpy roads that are less damaging to the vehicle.[18] We find that the distributions of lateral and vertical acceleration in the treatment group tightens around zero, consistent with a reduction in reckless and damaging

sex, owner education, owner self-employment status, the number of other businesses the owner runs, owner years of matatu industry experience, and owner raven score.

[16] It should be noted that this specification differs from the pre-specified regression, which pools all months of our data together and masks important dynamics.

[17] The monitoring device powers on and off with the matatu, allowing us to track when the vehicle is operating.

[18] The largest repair costs owners frequently face are run-down break pads, damaged shock absorbers, and broken axles.

driving.

In line with model predictions, these changes in driver behavior translate into lower repair costs. Figure 6 shows that repair costs reported by treatment owners decline steadily relative to control owners. The magnitude of the effect is large: by month six daily repair costs are $2 (KES 200) lower for treatment owners. This represents a 46% decrease in daily repair costs. They represent a major business expense for owners, which makes the impact of the monitoring technology significant.

According to the model, the effect of treatment on revenue is ambiguous and depends on whether the effect of lower risk or higher effort dominates. Owners may be willing to accept lower revenue if the reduction in repair costs from less risk more than offsets the reduction in expected transfers from lower revenue. The top panel of Appendix Figure A.8 supports this prediction: the effect on revenue is close to zero, and may be declining slightly. Proposition 2 also states that if revenue falls under the newly contracted effort-risk bundle, the owner should set a lower target to compensate the driver. While the effect is not statistically significant, the bottom panel of Appendix Figure A.8 suggests the target does indeed decrease over time.

Finally, we investigate the impact of the monitoring device on two measures of overall firm performance: profits and fleet size. Figure 6 shows that daily profits rise by approximately 12% in month four and five ($4.40 per day) for treatment owners. Taking the average gains over the study period and extrapolating to the full year (assuming matatus operate 25 days a month), owners can expect a $1,200 increase in annual firm profits. The device cost $125 (including shipping to Kenya), an amount that could be recovered in less than three months.

These profit gains are in line with some of the more successful business training programs documented in the literature (Bloom et al., 2013; McKenzie and Woodruff, 2017; Berge et al., 2015; de Mel et al., 2014; Valdivia, 2015). It is worth mentioning, however, that this profit measure does not capture any additional gains from having to spend less time and effort monitoring the driver, nor does it account for any additional costs incurred from increases in the firing probability. We discuss this further in Section 6.4.

Perhaps in response to higher profits, at endline we find that treatment owners have 0.145 more vehicles in their fleet on average than control owners (Table 3), an 11 percent increase in fleet size. Table 3 also presents suggestive evidence that owners invested in the interior of their vehicles through the purchase of items such as higher quality seating, lighting, and sound systems. This is consistent with the idea that owner's value their businesses more, and that a lack of monitoring was constraining firm growth.

6 Structural Estimation

The previous section provides reduced-form evidence for the effect of monitoring on driver behavior, firm outcomes, and contract parameters. But it cannot speak to the impact of monitoring on welfare as defined in Proposition 2: that is, the sum of the changes in implied contract valuation by the principal and the agent. To quantify this welfare effect, we first demonstrate that the model provides a close fit to the data in the status quo environment without monitoring. We then estimate unobservable parameters of the structural model that allow us to quantify welfare under the status quo and after the introduction of monitoring. Specifically, we numerically solve the owner's maximization problem, which we simplify by applying the Lemma to Equation (1):

$$V = \max_{T \in \mathcal{Y}} \delta V + T - G\left(T|e,r\right)\left(1 + \frac{V}{U}\right)\left\{T - \mathbb{E}\left[y|e,r,y \leq T\right]\right\} - \mathbb{E}\left[c(r)|r\right] \tag{3}$$

where

$$U = \frac{\mathbb{E}[y|e,r] - T - \psi(e,r)}{1-\delta}$$

is the driver's contract valuation under the target contract. Equation (1) makes explicit that V is the maximal value of the owner's objective function. Hence, solving for V involves finding the fixed point of this equation.

To find the target T that maximizes owner utility numerically, we need to empirically specify five terms in Equation 3. Two are calibrated from the data: the distribution of revenue $G(\cdot)$ (which also affects $\mathbb{E}\left[y|e,r,y \leq T\right]$) and the expected repair costs $\mathbb{E}\left[c(r)|r\right]$. Three are unobservable parameters that need to be estimated: the driver's outside option $\bar{u}$, firing costs h, and driver disutility from effort and risk $\psi(e,r)$ evaluated at (e_B, r_B) and (e_M, r_M).[19] We use simulated method of moments (SMM) as described below to simultaneously pin down these parameters and solve for the optimal target. We then use this output to evaluate the owner's contract valuation V specified in Equation (3).

Armed with these estimates of structural parameters under the status quo, we can then turn to evaluating the welfare effect of monitoring. We update our calibrations of the expected costs $\mathbb{E}[c(r)|r]$ and the revenue distribution $G(\cdot)$ using the observed treatment effects, and re-run the estimator with these new inputs. This allows us to estimate the change in contract valuation for both the owner and the driver.

[19] Because we evaluate $\psi(e,r)$ only at these two points – the effort-risk bundle at baseline (e_B, r_B) and under monitoring (e_M, r_M) – we do not need to make any functional form assumptions to estimate.

6.1 Simulated Method of Moments Estimation

We now describe how SMM allows us to estimate the unobserved parameters $\theta \equiv (\psi(e_B, r_B), h, \bar{u})$. SMM searches for the unknown parameters to minimize the weighted distance between a subset of the contract outcomes we can observe in the data and their simulated counterparts. The SMM estimator is:

$$\widehat{\theta} = \arg\min_{\theta} \left[\widehat{m}_D - \widehat{m}_S(\theta)\right]' W \left[\widehat{m}_D - \widehat{m}_S(\theta)\right]$$

where $\widehat{m}_D$ is the vector of moments calculated from the data, $\widehat{m}_S(\theta)$ is the analogous vector of endogenous moments simulated from the model for any given choice of θ, and W is the weighting matrix.[20] We include two different sets of moments for the two exercises below (i.e. baseline contract evaluation and monitoring valuation). To study the baseline contract, we use five moments

$$\widehat{m} = (\widehat{T}, \widehat{\mathbb{E}}[y - t|e, r], \widehat{\mathbb{E}}[t - c(r)|e, r], \widehat{\mathbb{E}}[p(t)|e, r], \widehat{U})'$$

using either estimated means from the data for $\widehat{m}_D$ or simulated means based on parameter guesses for $\widehat{m}_S(\theta)$.[21] For the first moment, $\widehat{T}_D$ is the average target observed in the data, while $\widehat{T}_S(\theta)$ is the numerically optimal target in Equation 3 for a given θ and a "bootstrapped" distribution of revenue $\tilde{G}(\cdot)$. Similarly, the next three moments use the average driver residual claim, the owner's net transfer, and the rehiring probability on the data side, and simulated counterparts using θ and $\tilde{G}(\cdot)$ on the simulated side. The final moment is driver welfare, for which we use the average reported price to forego the current contract as the data moment and the model expression for driver welfare evaluated at θ as the simulated moment.

In the second exercise (i.e. valuing the monitoring effect) we use a single moment: the change in the target $\widehat{T}$. We only require a single moment because we keep the outside option ($\bar{u}$) and firing costs (h) from the status-quo estimation fixed.[22] For both exercises, we run this simulation 500 times, where each simulation samples with replacement from the observed empirical distribution of revenue to create the bootstrapped distribution $\tilde{G}(\cdot)$.

6.2 Baseline Contract Valuation under Status Quo

Our first exercise estimates the baseline model parameters using data from the control group. Expected daily repair costs are fixed at \$4.99 (the mean in the control group) and the CDF of

[20] We use the optimal weighting matrix given by the inverse variance-covariance matrix of the data moments.

[21] In principle it would possible to estimate the parameters using only three moments, however using all five moments allows us to conduct an overidentification test of the model.

[22] Appendix Table A.3 shows robustness of our estimation to matching on treatment effects for salary and profits in addition to the target.

revenue $G(y|e,r)$ is sampled from the underlying data from the control group. We also assume a common discount factor of 0.99.

Table 4 summarizes the results of the status-quo model estimation. Starting with the parameter estimates, driver daily disutility of baseline effort and risk choices is estimated to be equivalent to approximately \$5, while firing costs are estimated to be equivalent to 51 days of lost profits. Both of these values are reasonable, as firing costs likely reflect a combination of idle days for the vehicle and the cost of finding a new driver. The driver's outside option is estimated to be \$7.07, which is approximately equal to the average unskilled daily wage in our context.[23]

The model also does a good job of matching the predicted moments from the simulation and the observed moments in the data. The target is matched nearly exactly, while expected profits and driver welfare are well within one standard deviation of what we observe in the data. Predicted driver salary is \$0.69 below observed salary, a relatively small difference. The only moment which the model fails to match by a meaningful margin is the firing probability, where the model predicts 0.6% of days will result in a separation versus the observed 0.1%.[24]

Finally, we can examine owner and driver welfare under the status quo. The model shows that the contract confers substantial welfare to the owner and the driver. Drivers receive \$442 above their outside option while owners benefit \$1,116 from this relationship with the driver. While drivers do not reach the target every day, their residual claim on days they do is substantially higher than their outside option. Owners have high repair costs but induce large transfers even when drivers miss the target.

6.3 Valuation with Monitoring

Our second exercise aims to estimate how the introduction of monitoring affects owner and driver welfare. We now rely on data from the treatment group to calculate expected repair costs and the CDF of revenue. We hold driver's outside option and firing costs fixed from the status quo estimation as these parameters are not expected to change under monitoring. The final unknown model parameter is driver disutility in the treatment group. The SMM procedure first calculates new model outcomes under monitoring, which determine a simulated treatment effect (by comparing the new simulated moments to those estimated under the status quo). The procedure then estimates what value of driver disutility produces a simulated treatment effect that best matches the observed treatment effect in the data.

Table 5 displays the results of the second SMM estimation. We find that driver disutility is

[23] Appendix Figures A.9 to A.12 show how sensitive the model outcomes are to different values of the model parameters. These Figures show that our calibration is robust to other reasonable parameter choices.

[24] To assess the overall fit of the model, we use an over-identification test of the model's overall fit. We find that we cannot reject the model (p=0.54).

estimated to have risen by 22% to \$6.08 relative to the status quo. If the disutility function is convex, then this effect is consistent with the reduced-form increases in both mileage (11%) and driving time (10%). The fit of the simulated moments of the treatment effects is somewhat lower than what we observe under the status-quo. The target treatment effect match is nearly perfect (although this is expected given that the estimation was exactly identified using this moment), while the treatment effect on the firing probability is also close. The predicted profit treatment effect (\$1.18 increase) is smaller than the observed profit treatment (\$3.62) though not statistically significant, while the predicted salary treatment effect is higher (\$0.93 versus \$0.23). That is, our model slightly underestimates the benefits to the driver and overestimates the benefits to the driver. We discuss this finding more below.

Most interestingly, this exercise estimates the changes in driver and owner welfare. From the theoretical model, we predict that owner welfare will rise after the introduction of monitoring. In contrast, the effect of monitoring on driver welfare and total welfare is ambiguous. It depends on how driver disutility changes as the status quo bundle of effort and risk shifts from (e_B, r_B) to (e_M, r_M) under monitoring. When driver disutility increases only slightly or falls, driver welfare rises. This outcome was not possible without monitoring because it could only be achieved by the owner setting a lower target, and the driver committing to a more favorable effort-risk bundle. Committing to such a bundle was not credible in the absence of monitoring.

Figure 7 Panel (b) illustrates these dynamics for different costs. For increases in driver disutility below 10% the driver is better off with monitoring than without. As driver disutility increases between 10% and 25%, monitoring lowers driver welfare but still raises overall welfare as the gains to the owner are larger than the losses to the driver. However, once driver disutility increases beyond 25%, driver losses dominate the owner's gains causing overall welfare to fall. Our SMM procedure estimates driver disutility increasing by 22%, which means driver welfare falls by \$31 while owner value increases by \$41, leading to a small overall welfare improvement of \$10. It is worth highlighting that there is some uncertainty on the overall welfare impacts, though it does seem that in this case the owner benefits more than the driver.[25]

6.4 Discussion of Structural Estimation Results

These structural estimation results suggest that monitoring leads to small efficiency gains, with some redistribution from the driver to the owner. We provide further context for these findings.

[25] If we estimate the model matching all three treatment effects, we find that both owner *and* driver valuation increase (see Appendix Table A.3). This result requires driver disutility to fall by 9% under the new bundle of effort and risk. This is consistent with the model's predictions, that allow for the possibility that both owner and driver welfare could increase with the introduction of monitoring. However, given the high number of working hours under the status-quo, it seems unlikely that driver disutility falls with the observed increase in effort.

First, to evaluate the plausibility of the owner's valuation of the monitoring device, we can compare our estimates to our findings from the willingness to pay experiment conducted at the end of the study. We estimate the owners' average willingness to pay for the monitoring system to be \$44.67 as compared to the model estimate of \$41.10. This suggests the model captures the benefits of monitoring to the owner well.

Second, we complement drivers' welfare estimates with SMS survey responses that we collected from drivers six months after the study finished.[26] Out of the 60% of drivers who responded, one quarter said the tracking device improved the relationship with their owner while nearly three quarters reported no change (only 3% reported a worse relationship). 96% said they preferred driving with the tracker. While this qualitative evidence may suffer from interviewer demand effects and selection, it does indicate that monitoring may have conferred non-pecuniary benefits to the driver. Consistent with this interpretation, we find that treatment owners transferred a slightly larger amount to their driver in a trust game at endline. They also believed their driver was more honest and his driving style has improved. We report these findings in Appendix Table A.2. This evidence suggests that an improvement in the owner-driver relationship may have counteracted some of the costs drivers' incurred from monitoring.

Finally, we discuss the fact that the model slightly underestimates profits but overestimates the driver's residual salary. One explanation is that the monitoring device may provide owners a weak signal of revenue, making it harder for drivers to withhold revenue when they fall below the target. While treated owners are slightly more likely to report that they know the revenue the driver collected on a given day (Table 2), the effect is small and insignificant. Despite the absence of this effect, treated drivers may expect their owners to have a better sense of the revenue they earn. This could prompt them to transfer more to the owner, thereby lowering the driver's salary and increasing the owner's profits.

7 Conclusion

A firm's success rides heavily on the performance of its employees. It is therefore important that firms design employment contracts that properly incentivize hard work. This becomes more challenging when firms cannot observe the amount of effort employees invest, nor the amount of output they produce. In theory, firms can use monitoring technologies that reveal the performance of their workers more accurately to overcome this constraint. In practice, however, the impact of such monitoring technologies on contracts and employee performance is unclear.

[26] At this point we had given control owners two months with app access as well, so no distinction can be drawn between treatment and control drivers. Response frequency was balanced across treatment and control drivers.

In this paper, we investigate the impact of providing firms with monitoring devices on contracts and firm productivity. This question is particularly relevant as information technologies are becoming ubiquitous. Our setting is one where the firm demands a fixed transfer, or "target" at the end of the day, which the agent pays from the revenues they collect. We develop a model that outlines the inefficiencies of this class of target contracts, which are prevalent in low-income countries. The model also generates predictions about the effect of monitoring on the principal, the agent, and firm outcomes. These predictions are then tested using a field experiment, where we provided a randomly selected sample of minibus owners in Kenya with a novel monitoring system we designed for the industry. The monitoring system reveals the amount of effort and risk the driver supplies, but is unable to provide information on the exact revenues they collect.

We find that the monitoring device allows minibus owners to demand a new bundle of effort and risk from the driver that was previously impossible to incentivize. This results in higher profits. Firms then reduce the transfer they demand from drivers to compensate them for the higher disutility they incur. We do not see the owners changing the contract structure, suggesting that target contracts are likely to remain widespread in this industry – at least until monitoring technologies can reveal the amount of revenue employees earn throughout the day.

We estimate the target contract model via simulated method of moments to determine whether the changes in contracts and productivity translate into welfare gains for the owner and driver. Both parties' value the contract considerably under the status quo, on the order of several months of average income. This may explain the prevalence of target contracts and the relative scarcity of owners operating the vehicles for themselves. We find that owners welfare increases by approximately $41 with the introduction of monitoring. While our model predicts that drivers could be better off under monitoring, our setting is one where the disutility from new effort and risk bundle outweighs the gains from a lower target. It is worth highlighting that these losses may be compensated by greater trust between owners and drivers.

Taken together, these results provide compelling evidence that this popular class of contracts are inefficient, and that monitoring devices can help firms overcome these inefficiencies. While our results shed light on the role of monitoring in contracting outcomes, they are also important from a policy perspective. They are particularly relevant for small firms, and policy makers working to improve the efficiency of the transportation industry. We know firms struggle to grow in developing countries for a number of reasons, and this paper identifies another important barrier that is relatively understudied empirically: moral hazard in labor contracting. We demonstrate that introducing cost-effective monitoring technologies can be a worthwhile investment for companies looking to expand. Similarly, we show that monitoring technologies can improve the operational efficiency of the informal transit industry provided information is conveyed to the owner. Currently

local transport authorities in different countries have started to discuss ways of introducing remote tracking solutions throughout the industry to help monitor and record the behavior of the drivers on the road.[27] While these policies focus primarily on safety considerations, our results demonstrate how these technologies could improve operational efficiencies as well.

We see a number of priorities for future research. First, this study focuses on a relatively small number of treatment owners spread across nine different routes. While this removes any concerns of spillovers within a route, it does mean that the impact on productivity and welfare could change as the proportion of owners with access to the technology increases. Second, our study focuses on a common monitoring device that provides information on key aspects of driver behavior. While this is currently the most relevant device in many markets, there are others systems such as digital payments devices that could change the contract structure altogether because they reveal information about revenue. Studying the impact of more sophisticated devices, at a larger scale could provide interesting avenues for future research. Finally, it is important to investigate how the entire regulatory environment could be improved, and whether the impact of the monitoring systems is greater under these circumstances.

[27] Kenya and South Africa have started engaging in these discussions. We investigate the effect of monitoring technologies on road safety in a companion paper

References

Baker, George P. and Thomas N. Hubbard (2004): "Contractibility and Asset Ownership: On-Board Computers and Governance in U. S. Trucking," *The Quarterly Journal of Economics*, Vol. 119, pp. 1443–1479.

Behrens, Roger, Dorothy McCormick, and David Mfinanga (2015): *Paratransit in African Cities: Operations, Regulation and Reform*: Routledge.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden (2015): "Human and Financial Capital for Microenterprise Development: Evidence from a Field and Lab Experiment," *Management Science*, Vol. 61, pp. 707–722.

Björkman, Martina and Jakob Svensson (2009): "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda," *The Quarterly Journal of Economics*, Vol. 124, pp. 735–769.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013): "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, Vol. 128, pp. 1–51.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts (2010): "Why Do Firms in Developing Countries Have Low Productivity?" *American Economic Review*, Vol. 100, pp. 619–23.

Bruun, Eric and Roger Behrens (2014): "Paratransit in Sub-Saharan African Cities—Improving and Integrating "Informal" Services," in *Shaping the New Future of Paratransit: An International Conference on Demand Responsive Transit, Monterey, CA, USA, October 29-31, 2014*, p. 18: Transportation Research Board.

Cervero, Robert and Aaron Golub (2007): "Informal transport: A Global Perspective," *Transport Policy*, Vol. 14, pp. 445 – 457.

de Mel, Suresh, David McKenzie, and Christopher Woodruff (2014): "Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka," *Journal of Development Economics*, Vol. 106, pp. 199 – 210.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan (2013): "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*," *The Quarterly Journal of Economics*, Vol. 128, pp. 1499–1545.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan (2012): "Incentives Work: Getting Teachers to Come to School," *American Economic Review*, Vol. 102, pp. 1241–78.

Ghatak, Maitreesh and Priyanka Pandey (2000): "Contract choice in agriculture with joint moral hazard in effort and risk," *Journal of Development Economics*, Vol. 63, pp. 303 – 326.

Godard, X (2006): "Coping with Paratransit in Developing Cities, a Scheme of Complementarity with Institutional Transport," in *Presentation to Future Urban Transport Conference, Göteborg, Sweden*, pp. 2–5.

Grossman, Sanford J. and Oliver D. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, Vol. 51, pp. 7–45.

Habyarimana, James and William Jack (2011): "Heckle and Chide: Results of a randomized road safety intervention in Kenya," *Journal of Public Economics*, Vol. 95, pp. 1438 – 1446, Special Issue: International Seminar for Public Economics on Normative Tax Theory.

——— (2015): "Results of a large-scale randomized behavior change intervention on road safety in Kenya," *Proceedings of the National Academy of Sciences*, Vol. 112, pp. E4661–E4670.

Hart, Oliver D and Bengt Holmström (1986): "The Theory of Contracts."

Holmström, Bengt (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, Vol. 10, pp. 74–91.

Hubbard, Thomas N. (2000): "The Demand for Monitoring Technologies: The Case of Trucking," *The Quarterly Journal of Economics*, Vol. 115, pp. 533–560.

——— (2003): "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking," *American Economic Review*, Vol. 93, pp. 1328–1353.

Innes, Robert D (1990): "Limited liability and incentive contracting with ex-ante action choices," *Journal of Economic Theory*, Vol. 52, pp. 45 – 67.

Kelley, Erin, Gregory Lane, and David Schönholzer (2020): "Agency and Improving Safety in Public Transit," *Working Paper*.

Kenya Roads Board (2007): "Kenyan Transport Sector Details Annex 3.1.,"Technical report, Tech. rep., Kenya Roads Board.

Levin, Jonathan (2003): "Relational Incentive Contracts," *American Economic Review*, Vol. 93, pp. 835–857.

Macchiavello, Rocco and Ameet Morjaria (2015): "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review*, Vol. 105, pp. 2911–45.

McCormick, Dorothy, Winnie Mitullah, Preston Chitere, Risper Orero, and Marilyn Ommeh (2013): "Paratransit Business Strategies: a Bird's-Eye View of Matatus in Nairobi," *Journal of Public Transportation*, Vol. 16, p. 7.

McKenzie, David and Christopher Woodruff (2017): "Business Practices in Small Firms in Developing Countries," *Management Science*, Vol. 63, pp. 2967–2981.

Mutongi, Kenda (2017): *Matatu: a History of Popular Transportation in Nairobi*: University of Chicago Press.

de Rochambeau (2020): "Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms," *Working Paper*.

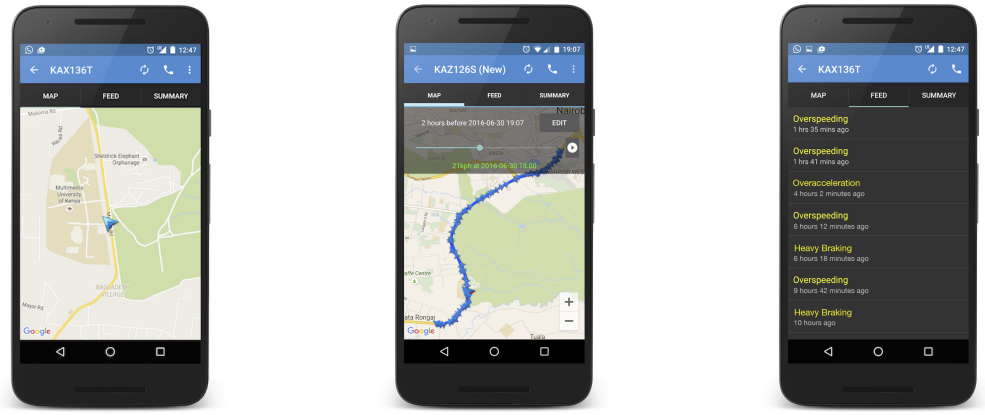
Townsend, Robert M (1979): "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, Vol. 21, pp. 265 – 293.

Valdivia, Martín (2015): "Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru," *Journal of Development Economics*, Vol. 113, pp. 33 – 51.

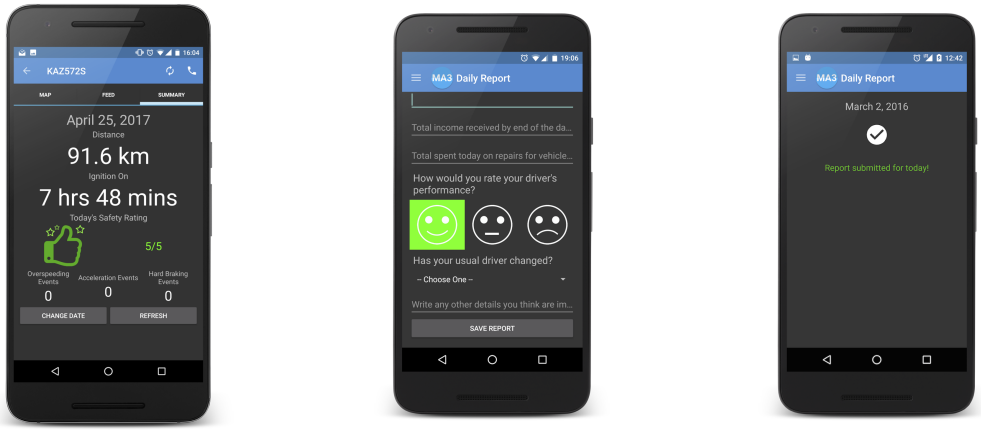
Figures

Figure 1: Mobile app "SmartMatatu"

(a) Map Viewer **(b)** Historical Map Viewer **(c)** Safety Feed



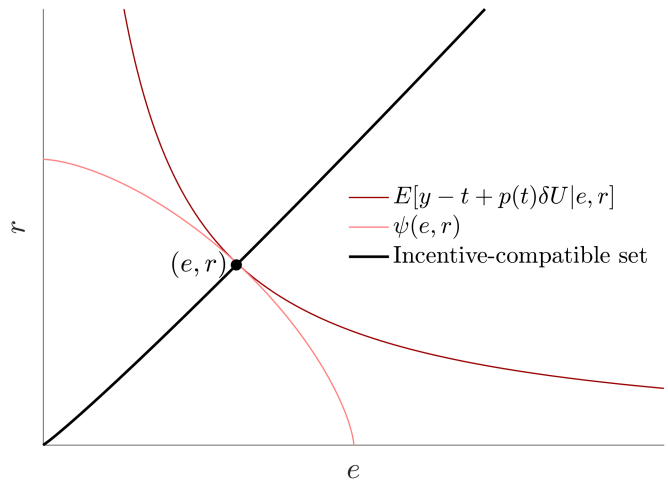
(d) Productivity Summary **(e)** Report Submit **(f)** Report Complete



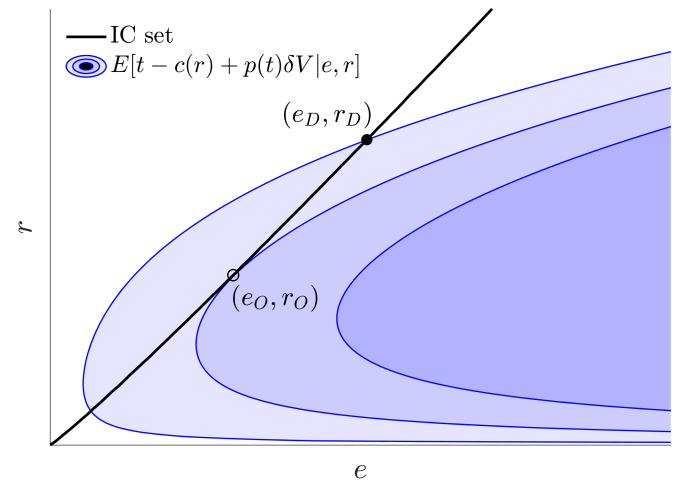
Notes: Android mobile app "SmartMatatu" developed by Echo Mobile in collaboration with matatu owners. Panels A and B: map viewer of real-time matatu location with historical playback of past locations over several hours for a given day. Panel C: Safety feed with speeding, acceleration, and hard braking alerts. Panel D: Daily productivity summary, with mileage in kilometers, number of hours ignition on as a measure of hours worked, and summary safety rating relative to other drivers on the route. Panels E and F: Reporting for both treatment and control owners of daily target, transfer received, repair costs, satisfaction with driver, and notification in case the driver changed.

Figure 2: Baseline and monitoring contract intuition

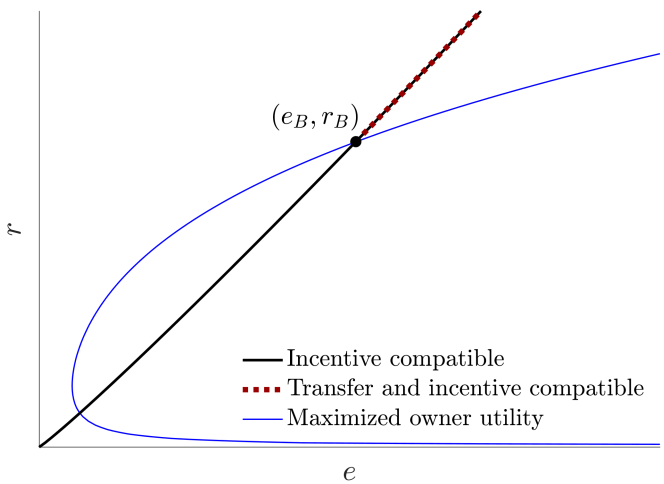
(a) Driver utility in (e, r) space.



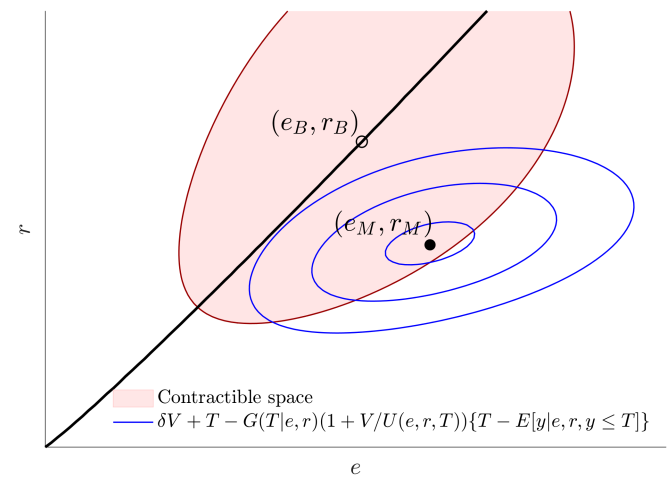
(b) Assumption 2 implies $(e_D, r_D) > (e_O, r_O)$.



(c) Incentive compatible contracting choice (e_B, r_B).

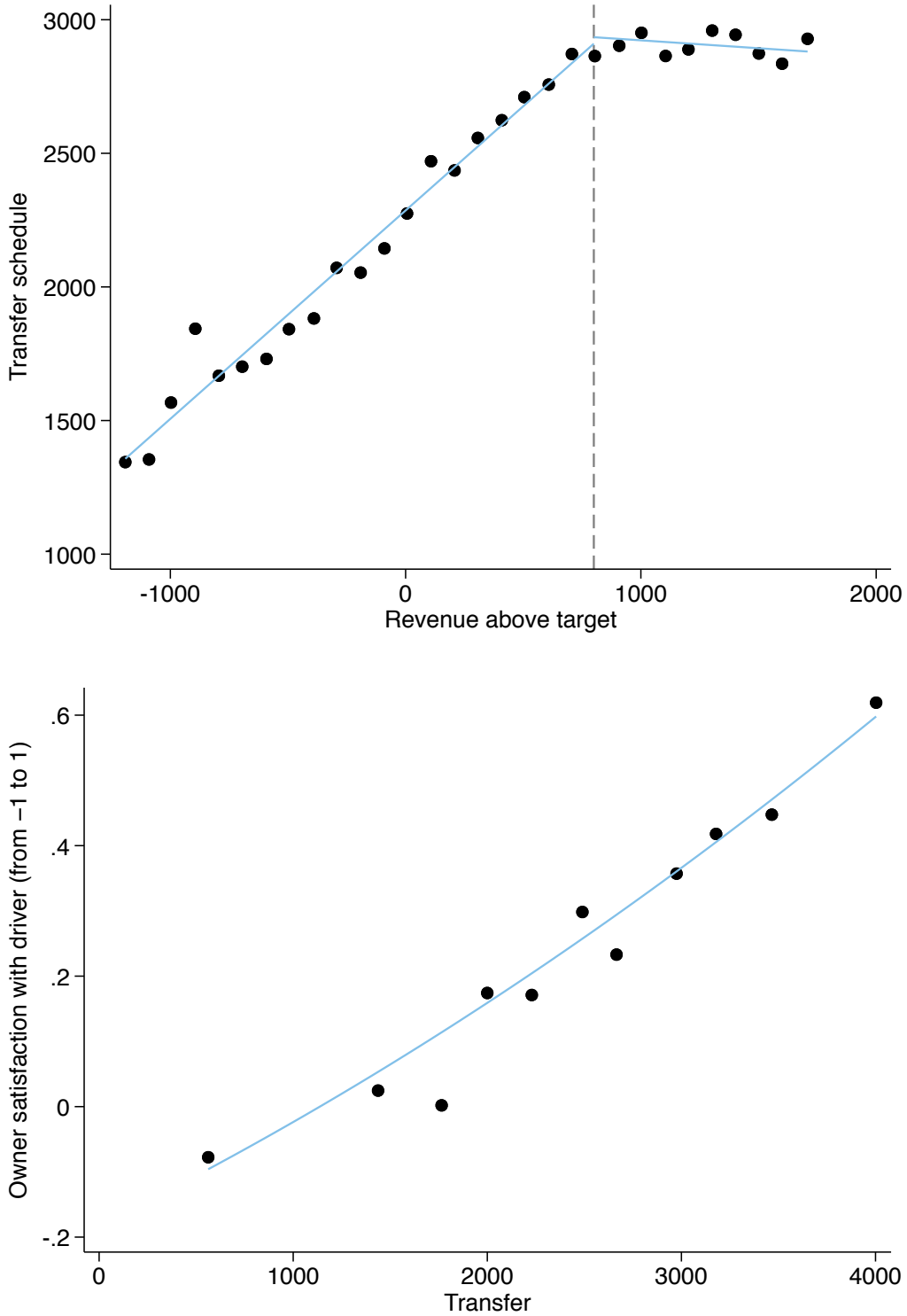


(d) Monitoring shifts effort/risk to (e_M, r_M).



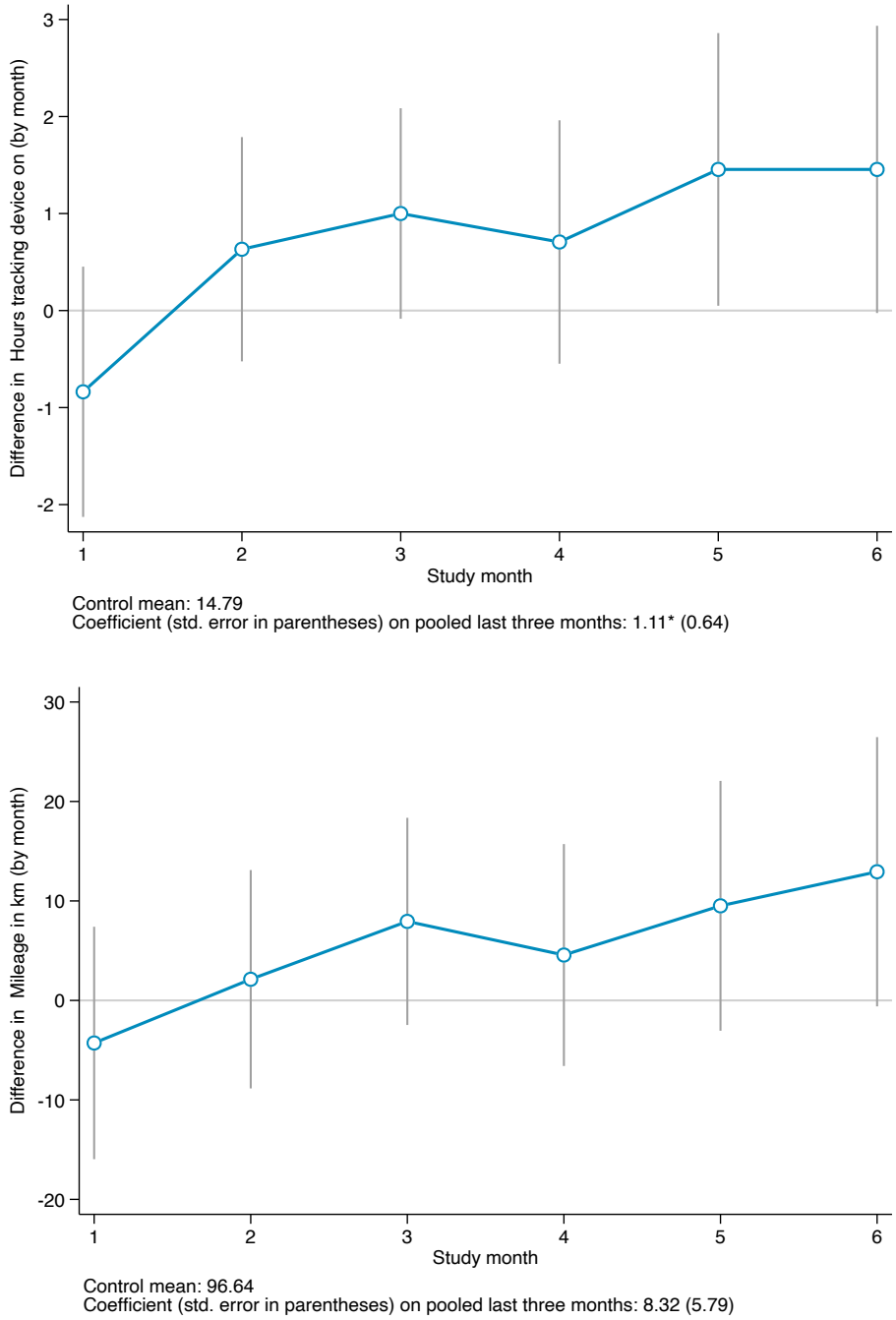
Notes: Panel A: The owner can only induce effort-risk bundles on the incentive compatible set. Panel B: The owner constrained bliss point (e_O, r_O) is below the driver bliss point (e_D, r_D) due to Assumption 2. Panel C: The baseline contracted bundle (e_B, r_B) coincides with the driver bliss point. Panel D: With Monitoring, effort rises and risk falls; the owner faces a tradeoff in effort and the target.

Figure 3: Estimated transfer schedule and owner satisfaction in response to transfer



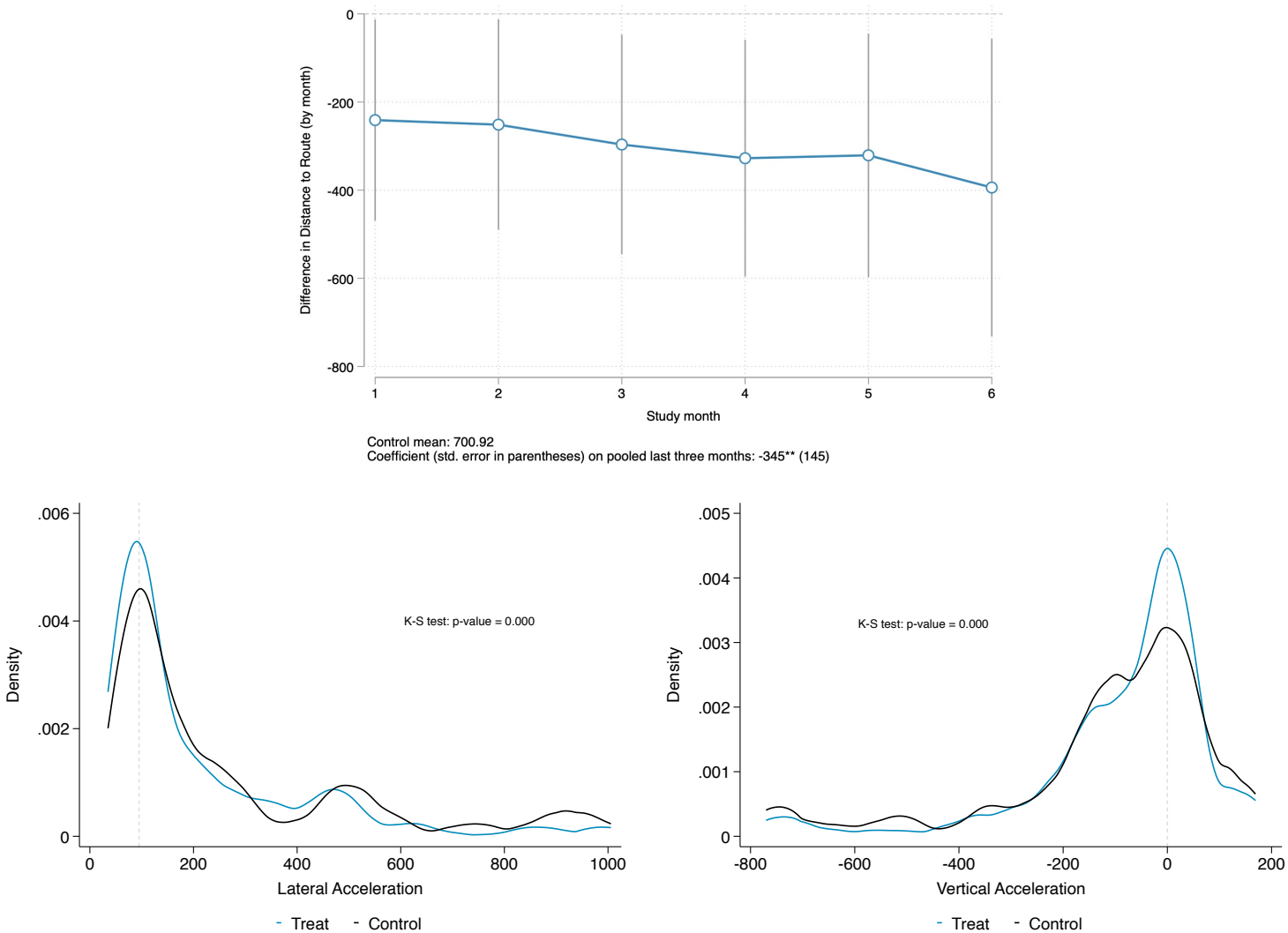
Notes: Top panel: The empirical transfer schedule closely resembles the shape $t(y) = \min\{y, T\}$ as in the Lemma. The slope extends beyond the target because of subsistence income, which we include in the structural estimation (see text). Bottom panel: Owner satisfaction rises substantially with the transfer, as suggested by $p(t) = p_0 + \frac{t}{\delta U}$. In Appendix Figure A.7 we also show that rehiring falls with transfers below the target.

Figure 4: Treatment effects on effort



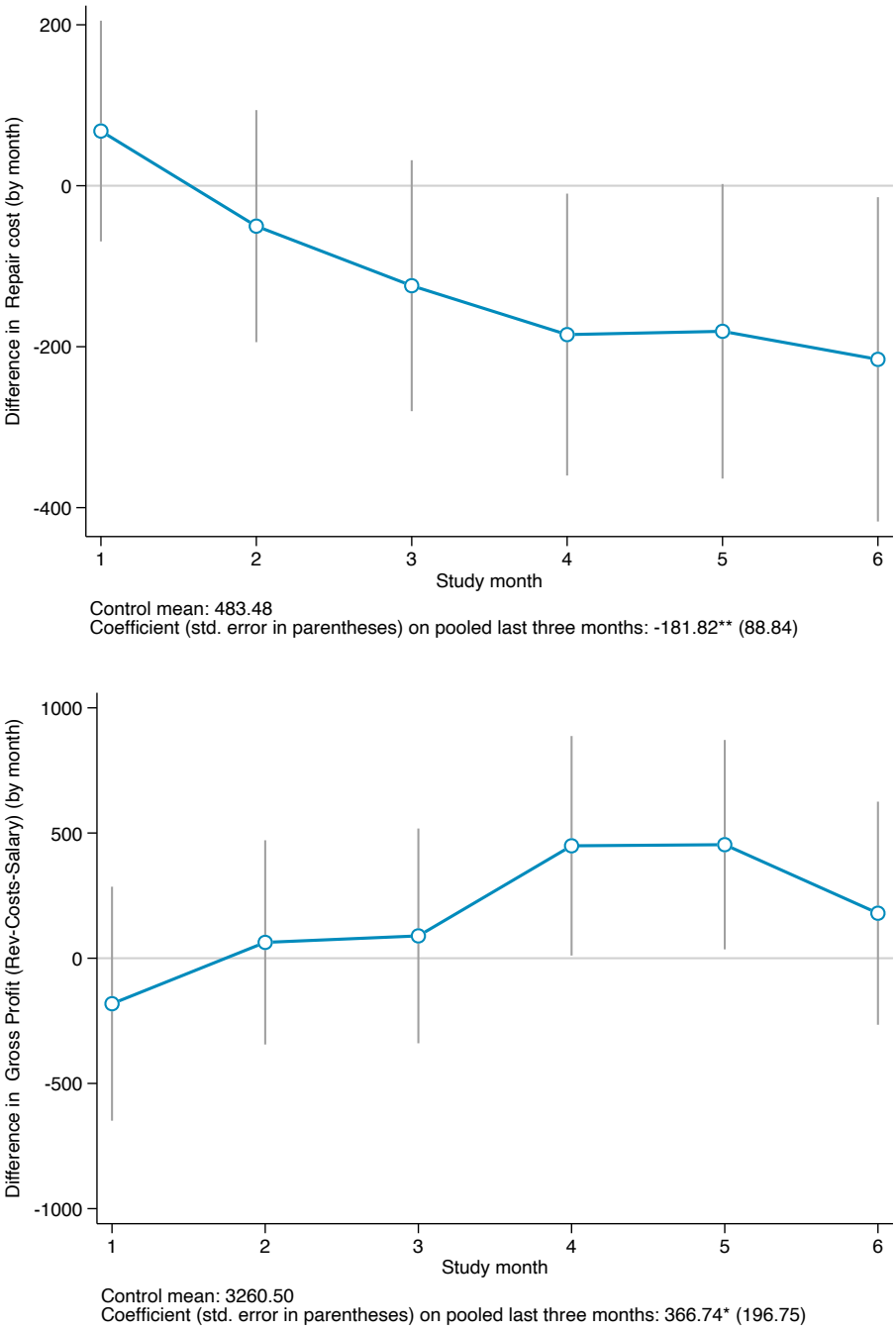
Notes: OLS estimates according to Equation 2. Top panel: Hours tracking device on corresponds to working hours of driver. Bottom panel: Daily mileage captured by tracking device. Standard errors for 95% confidence intervals clustered at the matatu level (255 clusters).

Figure 5: Treatment effects on risk taking



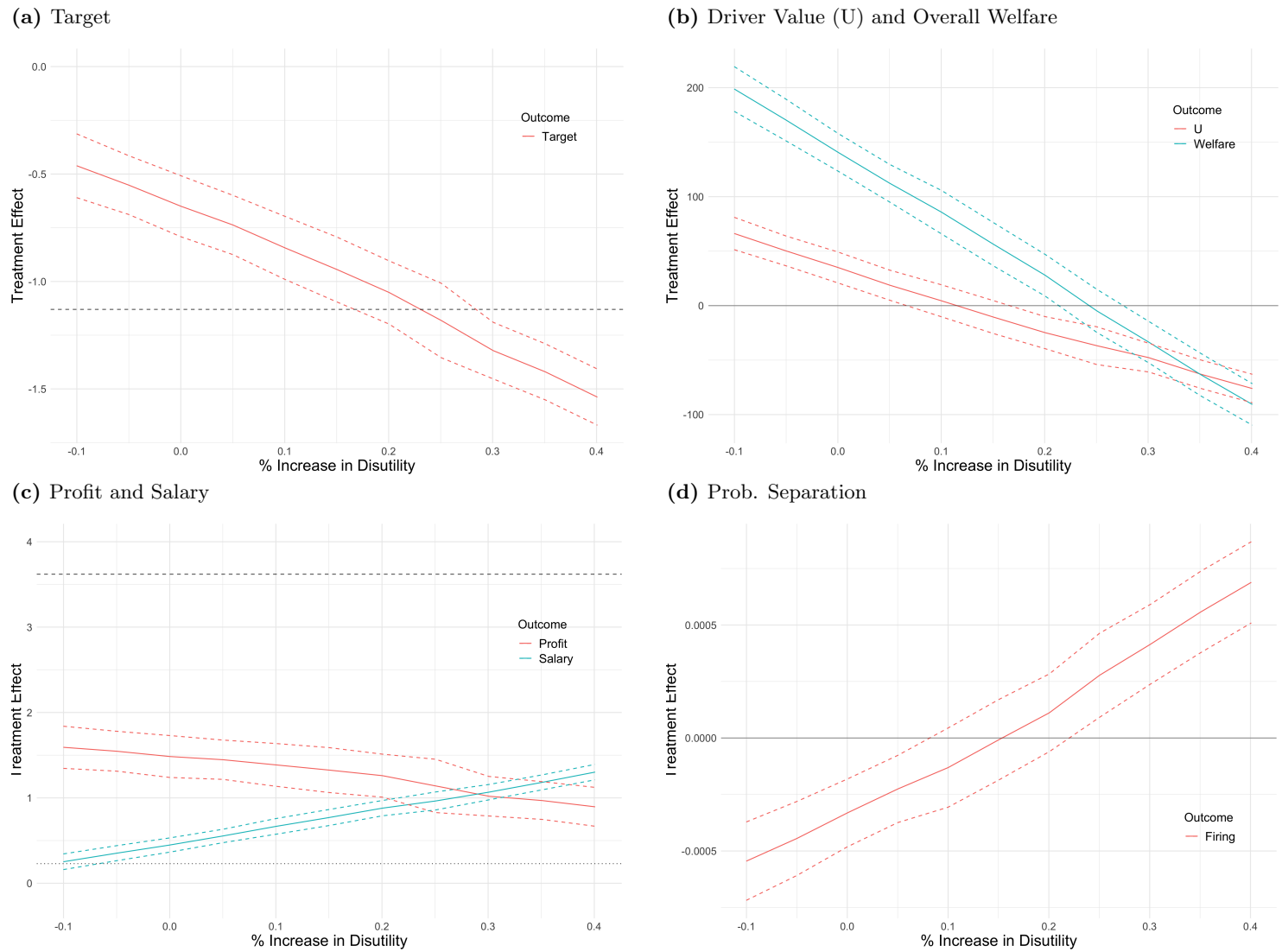
Notes: OLS estimates according to Equation 2. Top panel: Distance to licensed route in meters captured by tracking device. Middle panel: Treatment and control distributions of lateral acceleration. Standard errors for 95% confidence intervals clustered at the matatu level (255 clusters). Bottom panel: Treatment and control distributions of vertical acceleration. Kolmogorov-Smirnov test of equality of distributions.

Figure 6: Treatment effects on costs and profits



Notes: OLS estimates according to Equation 2. Top panel: Treatment effect by month on daily repair costs in KES as reported by owners. Bottom panel: Effect on gross profit, defined as revenue minus repair costs minus driver residual claim (salary). Standard errors for 95% confidence intervals clustered at the matatu level (255 clusters).

Figure 7: Predicted Treatment Effects by Increased Work Cost



Notes: Figures plot the model predicted treatment effect for different increases in driver work cost under monitoring. Dashed or dotted lines show observed values where applicable; in panel (c) the dashed line is the observed profit treatment effect while the dotted line is the observed salary treatment effect. Other input parameters are fixed per those stated in the main model calibration table.

Tables

Table 1: Summary statistics on owners, drivers, and matatus

	All				Treatment	Control	p-Value
	Mean	Std. Dev.	Min	Max	mean	mean	difference
Owners							
Age	36.72	7.87	18	68	37.16	36.29	0.377
Female	0.18	0.38	0	1	0.17	0.18	0.939
Years of education	11.65	2.88	0	14	11.52	11.77	0.501
Self-employed (yes/no)	0.78	0.41	0	1	0.77	0.79	0.689
Years industry experience	7.78	6.34	0	34	7.82	7.74	0.927
Years matatu owner	4.56	4.16	0	26	4.47	4.66	0.715
Number past drivers	1.85	1.73	0	10	1.94	1.77	0.437
Owner Raven's score	4.56	1.55	0	8	4.63	4.49	0.452
Owner rating: driver honesty	7.70	1.45	4	10	7.61	7.78	0.345
Owner rating: driver diligence	8.19	1.46	3	10	8.08	8.29	0.239
Baseline target	31.31	4.44	20	50	31.56	31.06	0.376
Baseline transfer	25.96	7.96	0	50	25.94	25.98	0.969
Drivers							
Age	35.71	7.25	21	58	37.19	34.27	0.001
Years of education	11.06	2.78	0	14	10.86	11.26	0.252
Years driving experience	7.89	5.89	0	37	8.75	7.05	0.021
Number of past owners	5.50	4.87	0	50	5.36	5.64	0.649
Months with current owner	14.77	19.90	0	180	14.25	15.27	0.684
Driver Raven's score	4.28	1.38	0	8	4.23	4.33	0.552
Driver risk choice	6.65	2.99	1	10	6.70	6.60	0.803
Driver rating: owner fairness	8.23	1.53	2	10	8.40	8.07	0.088
Baseline revenue	76.99	16.38	30	150	77.10	76.89	0.918
Baseline residual revenue	9.59	2.67	3	20	9.54	9.64	0.765
Matatus							
Age of matatu	13.06	4.27	2	26	13.46	12.67	0.142
Number of special features	1.38	0.89	1	8	1.40	1.37	0.825
Purchase price (USD)	6675	2849	1800	30000	6396	6947	0.123
Observations	255				126	129	

Notes: All values in 100s of Kenyan Shillings (KES, approximately $1). Data from baseline survey. Years of education constructed from categories, assuming partial completion (elementary: 4 years; high school: 10 years; university: 14 years; technical college: 12 years). Ratings of honesty and diligence (owner) and fairness (driver) range from 1 to 10. Driver risk choice based on a standard risk lottery game. p-value of t-test comparing means in treatment and control groups.

Table 2: Treatment effects on reported knowledge and monitoring behavior

	(1) Know mileage	(2) Know off-route	(3) Better driving	(4) Know revenue	(5) Difficulty monitor	(6) Monitoring time
Treatment	0.27***	0.45***	0.63***	0.04	-1.85***	-0.72***
	(0.07)	(0.07)	(0.06)	(0.07)	(0.16)	(0.05)
Control Mean of DV	0.47	0.40	0.04	0.61	4.02	-0.01
Controls	X	X	X	X	X	X
Route FE	X	X	X	X	X	X
Matatu N	187	187	190	187	190	190

Notes: OLS regressions of outcome on treatment indicator, controlling for route fixed effects, the age of the matatu, the number of special features, owner age and sex, owner education, owner self-employment status, the number of other businesses the owner runs, owner years of matatu industry experience, and owner raven score. "Know mileage": whether the owner knows the approximate number of kilometers a driver drove on a given day. "Know off-route": the owner knows when the driver is off the licensed route. "Better driving": the owner's judgement of overall driver performance at endline, worse (-1), about the same (0), or better (1). "Know revenue": the owner know the approximate amount of revenue the driver made. "Difficulty monitor": how hard it is to monitor the driver's behavior, from 1 (very easy) to 5 (very hard). "Monitoring time": whether the owner's time spent monitoring the driver has increased (1), stayed the same (0), or fallen (-1) over the last six months. Data from additional question added to endline after one quarter of endlines were already completed, hence only up to 190 out of 255 observations (balanced across treatment and control; three non-responses to first three questions). Robust standard errors. Stars indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 3: Treatment effects on business investment

	(1) Number vehicles	(2) New interior
Treatment	0.129*	0.074
	(0.076)	(0.057)
Control Mean of DV	1.22	0.21
Controls	X	X
Route FE	X	X
Matatu N	245	240

Notes: OLS regressions of outcome on treatment indicator, controlling for route fixed effects, the age of the matatu, the number of special features, owner age and sex, owner education, owner self-employment status, the number of other businesses the owner runs, owner years of matatu industry experience, and owner raven score. "Know mileage" ask whether the owner knows the approximate number of kilometers a driver drove on a given day. "Know off-route": the owner knows when the driver is off the licensed route. "Know revenue": the owner know the approximate amount of revenue the driver made. "Number vehicles": the number of vehicles the owner owns at endline. "New interior": major investment into interior of vehicle. Data from endline survey. Robust standard errors. Stars indicate statistical significance at the 1% ***, 5% **, and 10% * levels..

Table 4: Model calibration under baseline contract

Panel A: Assumptions

Input	Symbol	Value	Justification
Repair Costs	$E[c(r)]$	4.9	Average repair costs seen in the control group
Revenue Distribution	$G(\cdot)$	—	Empirical distribution calculated from control group data
Discount Factor	δ	0.99	

Panel B: SMM Parameter Estimates

Input	Symbol	Value	Interpretation
Disutility	$\psi(e^*, r^*)$	4.99 (0.01)	Disutility equivalent to 5 dollars
Firing Cost	h	51.44 (0.50)	Total firing costs equivalent to 51 days lost profits
Outside Option	$\bar{u}$	7.07 (0.04)	Approximately the average unskilled daily wage

Panel C: Matched Moments and Predictions

Outcome	Prediction	Observed	Difference
Target	30.13	30.12	0
	(0.09)	(0.35)	(0.36)
Expected Profit	22.41	23.25	-0.843
	(0.14)	(0.89)	(0.9)
Expected Salary	19.26	19.94	-0.69
	(0.03)	(0.32)	(0.32)
Prob. Separation	0.0064	0.0011	0.0054
	(0.00011)	(0.0002)	(0.0002)
Driver Contract Value minus Outside Option	442.12	446.18	-4.06
	(2.53)	(184.62)	(184.64)
Owner Value	1116.27	—	—
	(15.03)		
Total Welfare	1558.39	—	—
	(15.44)		

Notes: Simulated method of moments (SMM) calibration for work cost, firing cost, and outside option. Matched on five moments: observed target, profits, salary, separation probability, and driver contract value. All values are in 100s of Kenyan Schillings. Observed salary includes estimated fare collector salary.

Table 5: Reduced-form versus structural treatment effect estimation

Panel A: Assumptions

Input	Symbol	Value	Justification
Repair Costs	$E[c(r)]$	2.7	Previous value of 4.9 minus observed treatment effect reduction of 2.2
Revenue Distribution	$G(\cdot)$	—	Empirical distribution calculated from treatment group data
Discount Factor	δ	0.99	
Firing Cost	h	51.44	Fixed from previous value
Outside Option	$\bar{u}$	7.07	Fixed from previous value

Panel B: SMM Parameter Estimates

Input	Symbol	Value	Interpretation
Disutility	$\psi(e^*, r^*)$	6.08	Increase of 1.09 (22%) from previous value
		(0.07)	

Panel C: Predicted vs. Observed Treatment Effects

Outcome	Pred.Treat	Obs.Treat	Difference
Target	-1.14	-1.13	-0.01
	(0.14)	(0.87)	(0.88)
Expected Profit	1.18	3.62	-2.44
	(0.2)	(1.96)	(1.97)
Expected Salary	0.93	0.23	0.7
	(0.03)	(0.35)	(0.35)
Prob. Separation	0.0002	-0.01	0.01
	(0.00016)	(0.01)	(0.01)
Driver Contract Value minus Outside Option	-30.98	—	—
	(3.38)		
Owner Value	41.1	—	—
	(22.25)		
Total Welfare	10.12	—	—
	(22.65)		

Notes: All values are in 100s of Kenyan Schillings. Work cost estimated via simulated method of moments matching the observed target treatment effect.

Online Appendix for "Monitoring in Target Contracts: Theory and Experiment in Kenyan Public Transit"

Erin Kelley, Gregory Lane, and David Schönholzer

Contents

A Proofs

A.1 Proof of Lemma

Proof. The proof proceeds in two steps: in step 1, we show that under Assumption 1 the minimally optimal slope of $p(t)$ is $\frac{t}{\delta U}$ – that is, a $p(t)$ with lower slope than $\frac{1}{\delta U}$ cannot be optimal. In Step 2, we then show that under Assumption 2 this minimal slope is preferred to higher slopes.

We begin with Step 1. Define the target $T = \min\{t \in \mathcal{Y} : p(t) = 1\}$. For this to exist, we need that there exists a t for which $p(t) = 1$. Suppose this were not the case so that the optimal $p(t) < 1$ for all t. Then, in particular, $p(\bar{y}) < 1$. However, the owner would be strictly better off by setting $p(\bar{y}) = 1$: she would capture a higher continuation value without lowering transfer or driving incentives for the driver. Hence, $p(t) = 1$ for some $t \leq \bar{y}$ and T exists.

We next show that the minimal slope is necessary to induce maximal transfers for any given realization of $y \in \mathcal{Y}$. According to the LLC, $t(y) \leq y$. Note that whenever $t(y) = y$, it has to hold that for any $t, t' \in [0, T]$ and $t > t'$, the driver always transfers the larger amount t if $p(t)$ satisfies the following condition:

$$
\begin{aligned}
y - t + p(t)\delta U &\geq y - t' + p(t')\delta U \\
\frac{p(t) - p(t')}{t - t'} &\geq \frac{1}{\delta U}
\end{aligned}
$$

and in this case $t(y) = \min\{y, T\}$.

We can now show that there is no way to lower effort and risk incentives below the minimal slope. To this end, define the "transfer set" $\mathcal{T} = \{y \in [0, T] : t(y) = y\}$ to be all revenue realizations for which the rehiring schedule $p(\cdot)$ induces transferring all revenue. Hence, $p(y)\delta U \geq y$ whenever $y \in \mathcal{T}$. Let the complement to the transfer set be $\mathcal{Y} \backslash \mathcal{T} = \bigcup_{i=1}^{I} \mathcal{X}_i$ where $\mathcal{X}_i = (t_i, x_i]$ are connected sets with lower bound $t_i = \max\{y \in \mathcal{T} : y < x_i\}$ and $t(y) = t_i < y$ whenever $y \in \mathcal{X}_i$. Revenue realizations that fall into an interval $\mathcal{X}_i$ trigger transfers at the lower bound of $\mathcal{X}_i$ because all revenue beyond this lower bound has a higher direct return to the driver than its return in terms of increased future discounted contract value $p(t)\delta U$. We now split the owner's objective function

$\mathbb{E}\left[y - t + p(t)\delta U|e,r\right]$ into intervals $\mathcal{T}$ and $\mathcal{X}_i$ for $i = 1, ..., I$:

$$\begin{aligned}
\mathbb{E}\left[y - t + p(t)\delta U|e,r\right] &= \mathbb{E}\left[p(y)\delta U|e,r,y \in \mathcal{T}\right]\Pr\left(y \in \mathcal{T}\right) \\
&\quad + \sum_{i=1}^{I} \mathbb{E}\left[y - t_i + p(t_i)\delta U|e,r,y \in \mathcal{X}_i\right]\Pr\left(y \in \mathcal{X}_i\right) \\
&= \int_{y\in\mathcal{T}} \delta U p(y) g(y|e,r)dy \\
&\quad + \sum_{i=1}^{I} \left\{\left[p(t_i)\delta U - t_i\right]\Pr\left(y \in \mathcal{X}_i\right) + \int_{y\in\mathcal{X}_i} y g(y|e,r)dy\right\}.
\end{aligned}$$

The marginal effect of increasing $s \in \{e, r\}$ on the owner's utility is then

$$\int_{y\in\mathcal{T}} \delta U p(y) g_s(y|e,r)dy + \sum_{i=1}^{I} \int_{y\in\mathcal{X}_i} y g_s(y|e,r)dy.$$

All that is left to do in Step 1 is to show that this marginal effect is bounded from below by application of the minimal slope of $p(\cdot)$. Since $p(\cdot)$ only appears in the first term (i.e. those in the transfer set), we can ignore the second (i.e. the one with the non-transfer sets $\mathcal{X}_i$). According to the definition of $\mathcal{T}$, $p(y)\delta U \geq y$. Together with the MLRP, this implies that

$$\int_{y\in\mathcal{T}} \delta U p(y) g_s(y|e,r)dy \geq \int_{y\in\mathcal{T}} y g_s(y|e,r)dy,$$

meaning that there is no way to incentivize less effort and/or risk with any choice of $p(\cdot)$: marginal incentives are bounded from below by $\int_{y\in\mathcal{Y}} y g_s(y|e,r)dy$.

We now move to Step 2: that under Assumption 2, the owner never benefits from inducing higher effort or risk with a steeper rehiring schedule $p(t)$. To see this, write owner utility under at least minimal slope (i.e. with $p(t) \geq p_0 + \frac{t}{\delta U}$) as:

$$X(e,r) = \mathbb{E}\left[t - c(r) + p(t)\delta V|e,r\right] \tag{4}$$

$$= \int_0^T \left[y + p(y)\delta V\right] g\left(y|e,r\right) dy + \left(1 - G\left(T|e,r\right)\right)\left[T + \delta V\right] - \mathbb{E}\left[c(r)|r\right] \tag{5}$$

and the corresponding marginal effect of effort and risk:

$$X_s\left(e,r\right) = \int_0^T \left[y + p(y)\delta V\right] g_s\left(y|e,r\right) dy - G_s\left(T|e,r\right)\left[T + \delta V\right] - \frac{\partial \mathbb{E}\left[c(r)|r\right]}{\partial s}.$$

It remains to be shown that $\sum_{s\in\{e,r\}} \psi_s\left(e,r\right) X_s\left(e,r\right) \leq 0$ for all $(e,r) \geq (e_D, r_D)$. If this inequality

holds, then the owner's marginal utility in the direction of the driver's disutility gradient is negative: as the driver exerts more effort and risk past his bliss point (e_D, r_D) in the preferred direction of the driver, the owner's utility falls.

We construct an upper bound of this marginal effect by setting (a) $T = \bar{y}$, (b) $p(y) = 0$ if $y \leq y^*$ and 1 otherwise, where y^* is defined as the smallest y for which $g_s(y|e,r) \leq 0$ for all $y \leq y^*$, and (c) $V = \frac{\mathbb{E}[y|e,r]}{1-\delta}$. In this way, the owner captures all of the marginal benefit of effort and risk, which maximizes the returns to the owner. We then have

$$X_s(e,r) < \frac{\partial \mathbb{E}\left[(1+\phi_s)\, y - c(r)|e,r\right]}{\partial s}$$

with $\phi_s = -\frac{\delta}{1-\delta} G_s(y^*|e,r)$. The right-hand side expression is the result of applying the extreme conditions (a)-(c) from the last paragraph, where the owner receives maximal marginal returns to effort and risk. According to Assumption 2, we then have for all $(e,r) \geq (e_D, r_D)$:

$$\sum_{s \in \{e,r\}} \psi_s(e,r)\, X_s(e,r) < \sum_{s \in \{e,r\}} \psi_s(e,r) \frac{\partial \mathbb{E}\left[(1+\phi_s)\, y - c(r)|e,r\right]}{\partial s} \leq 0,$$

with the latter inequality holding because it corresponds to the owner's directional derivative falling as we move up the incentive compatible set.

□

A.2 Proof of Proposition 1

Proof. Using the Lemma, we can write the driver's problem as: $\delta U - T + \mathbb{E}[y|e,r] - \psi(e,r)$. Thus, the FOCs for the driver and the owner-driver are, respectively:

$$\frac{\partial \mathbb{E}[y|e^*,r^*]}{\partial r} - \frac{\partial \mathbb{E}[c(r^*)|r^*]}{\partial r} = \frac{\partial \psi(e^*,r^*)}{\partial r}$$
$$\frac{\partial \mathbb{E}[y|e_B,r_B]}{\partial r} = \frac{\partial \psi(e_B,r_B)}{\partial r},$$

while

$$\frac{\partial \mathbb{E}[y|e,r]}{\partial e} = \frac{\partial \psi(e,r)}{\partial e}$$

holds for both the driver and the owner-driver.

Write

$$\frac{\partial \mathbb{E}[y|e,r]}{\partial r} - \mu \frac{\partial \mathbb{E}[c(r)|r]}{\partial r} = \frac{\partial \psi(e,r)}{\partial r}$$

as the marginal problem that nests both the driver's and the owner-driver's optimal risk problem:

note that $\mu = 0$ is the driver's problem and $\mu = 1$ is the owner-driver's problem.

We now use the Implicit Function Theorem (IFT) to show that $\frac{\partial r}{\partial \mu} < 0$, which implies the first part of the statement, i.e. $r_B > r^*$.

Define $H : \mathcal{S} \to \mathbb{R}^2$ in the following way:

$$
\begin{aligned}
H_1(e,r) &= \frac{\partial \mathbb{E}[y|e,r]}{\partial e} - \frac{\partial \psi(e,r)}{\partial e} = 0 \\
H_2(e,r) &= \frac{\partial \mathbb{E}[y|e,r]}{\partial r} - \mu \frac{\partial \mathbb{E}[c(r)|r]}{\partial r} - \frac{\partial \psi(e,r)}{\partial r} = 0
\end{aligned}
$$

We can now apply the IFT:

$$
\begin{aligned}
\begin{bmatrix} \frac{\partial e}{\partial \mu} \\ \frac{\partial r}{\partial \mu} \end{bmatrix} &= - \begin{bmatrix} \frac{\partial H_1}{\partial e} & \frac{\partial H_2}{\partial e} \\ \frac{\partial H_1}{\partial r} & \frac{\partial H_2}{\partial r} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial H_1}{\partial \mu} \\ \frac{\partial H_2}{\partial \mu} \end{bmatrix} \\
&= \frac{1}{A} \begin{bmatrix} \left\{ \frac{\partial^2 \psi(e,r)}{\partial r \partial e} - \frac{\partial^2 E[y|e,r]}{\partial r \partial e} \right\} \frac{\partial E[c(r)|r]}{\partial r} \\ \left\{ \frac{\partial^2 E[y|e,r]}{\partial e^2} - \frac{\partial^2 \psi(e,r)}{\partial e^2} \right\} \frac{\partial E[c(r)|r]}{\partial r} \end{bmatrix}
\end{aligned}
$$

where $A = \frac{\partial H_1}{\partial e} \frac{\partial H_2}{\partial r} - \frac{\partial H_1}{\partial r} \frac{\partial H_2}{\partial e} > 0$. Thus, for $\frac{\partial r}{\partial \mu} < 0$ we need the following to be true:

$$
\begin{aligned}
\left\{ \frac{\partial^2 \mathbb{E}[y|e,r]}{\partial e^2} - \frac{\partial^2 \psi(e,r)}{\partial e^2} \right\} \frac{\partial \mathbb{E}[c(r)|r]}{\partial r} &< 0 \\
\Leftrightarrow \frac{\partial^2 \mathbb{E}[y|e,r]}{\partial e^2} &< \frac{\partial^2 \psi(e,r)}{\partial e^2}
\end{aligned}
$$

which holds according to Assumption 1. $\frac{\partial e}{\partial \mu} > 0$ holds in case

$$
\frac{\partial^2 \psi(e,r)}{\partial r \partial e} > \frac{\partial^2 \mathbb{E}[y|e,r]}{\partial r \partial e},
$$

which may or may not be true according to Assumption 1.

□

A.3 Proof of Proposition 2

Proof. To incentivize maximal transfers for any realization of y, the slope of the rehiring schedule continues to be bounded from below: that is, for any t, t' with $t > t'$ in $[0, T]$, where $T = \inf\{t : p(t,e,r) = 1\}$, it has to hold that

$$
\frac{p(t,e,r) - p(t',e,r)}{t - t'} \geq \frac{1}{\delta U}.
$$

The rehiring schedule also continues to be bounded from above by this slope. To see this, consider some rehiring schedule $\hat{p}(t,e,r)$ with higher than minimal slope. This implies there exists some $y \in \mathcal{Y}$ such that $\hat{p}(y,e,r) = 1$ but $p(0,e,r) + \frac{y}{\delta U} < 1$ and hence $\hat{T} = \inf\{\hat{p}(t,e,r) = 1\}$. For every such rehiring schedule, there exsists another with equal expected rehiring probability but at minimal slope, i.e. $\mathbb{E}\left[p_0 + \frac{t}{\delta U}|e,r\right] = \mathbb{E}\left[\hat{p}(t,e,r)|e,r\right]$ with $T = (1-p_0)\,\delta U > \hat{T}$. Because the target of the minimal slope rehiring schedule is strictly higher while the rehiring probability is the same, the owner will strictly prefer the minimal slope rehiring schedule.

To complete the argument for the functional form of the rehiring schedule, note that the owner can induce a particular (e_M, r_M) by setting $p\,(t,e,r)$ such that

$$\mathbb{E}\left[y - t + p(t,e_M,r_M)\delta U_M|e_M,r_M\right] - \psi\left(e_M,r_M\right) \geq \mathbb{E}\left[y - t + p\left(t,e,r\right)\delta U|e,r\right] - \psi\left(e,r\right)$$

for all $(e,r) \in \mathcal{S}$. If such a $p\,(t,e,r)$ exists, then setting $p\,(t,e,r) = 0$ for all $(e,r) \neq (e_M,r_M)$ satisfies this constraint. Existence of this sufficient $p(t,e,r)$ is guaranteed by the expected dynamic enforcement constraint $U_M \geq \mathbb{E}\left[y|e,r\right] - \psi\left(e,r\right)$. Putting together the bounds on the slope of the rehiring schedule and the condition on (e,r) that induce a positive rehiring probability, it follows that

$$p(t,e,r) = \begin{cases} p_M + \frac{t}{\delta U_M} & \text{if } e = e_M \text{ and } r = r_M \\ 0 & \text{otherwise} \end{cases}$$

is a solution to the problem.

For the second part of the Proposition, the owner problem with minimal slope is:

$$\max_{(e,r)\in\mathcal{S},T\in\mathcal{Y}} \delta V + T - G\left(T|e,r\right)\left(1 + \frac{V}{U(e,r,T)}\right)\left\{T - \mathbb{E}\left[y|e,r,y \leq T\right]\right\} - \mathbb{E}\left[c(r)|r\right]. \tag{6}$$

subject to the participation constraint and the expected dynamic enforcement constraint. To show that $e_M > e_B$ and $r_M < r_B$, we first show that the first derivatives with respect to e and r have the required sign at (e_B, r_B): owner utility rises with larger e and falls with larger r. We then argue that they continue to do so in the relevant subset of $\mathcal{S}$ until they either run up against a constraint or reach an interior solution by crossing zero. Partial derivatives with respect to $s \in \{e,r\}$ are:

$$\begin{aligned} &- G_s\left(T|e,r\right)\left(1 + \frac{V}{U(e,r,T)}\right)\left\{T - \mathbb{E}\left[y|e,r,y \leq T\right]\right\} \\ &+ G\left(T|e,r\right)\frac{V}{U(e,r,T)^2}\frac{\partial U}{\partial s}\left\{T - \mathbb{E}\left[y|e,r,y \leq T\right]\right\} \\ &+ G\left(T|e,r\right)\left(1 + \frac{V}{U(e,r,T)}\right)\frac{\partial \mathbb{E}\left[y|e,r,y \leq T\right]}{\partial s} - \frac{\partial \mathbb{E}\left[c(r)|r\right]}{\partial s}, \end{aligned}$$

where $U(e,r,T) = (\mathbb{E}[y|e,r] - \psi(e,r) - T)/(1-\delta)$. The first and the third additive term are always positive. The second term is zero at (e_B, r_B). The fourth term is always zero for e. Hence the partial with respect to e is positive at (e_B, r_B), as desired. For the partial with respect to r at (e_B, r_B) to be negative, we need the last term (i.e. expected marginal cost of risk) to outweigh the sum of the first and the third term. This is guaranteed by Assumption 2.

As we move into (e,r) with $e > e_B$ and $r < r_B$, the second term of the partial with respect to e becomes negative and grows at a faster rate than the first and the third term, guaranteeing that it eventually crosses zero. The second term of the partial with respect to r becomes positive for (e,r) with $e > e_B$ and $r < r_B$. But Assumption 2 guarantees that the partial as a whole remains negative for all $(e,r) \geq (e_B, r_B)$, and hence it will cross zero (or hit a constraint) at some $r < r_B$.

The result that profit increases follows directly from it being collinear with owner utility; (e_B, r_B) being in the owner's choice set; and owner utility increasing strictly when moving towards (e_M, r_M).

To see that revenue may rise or fall, note that (e_M, r_M) is in the set $\mathcal{S}_M = (e_B, \bar{e}] \times (0, r_B)$ and recall that $G_{e,r}(e,r) \geq 0$ from Assumption 1, which implies that the isoquant at (e_B, r_B) is downward sloping. Hence, the intersection of $\mathcal{S}_M$ with both the upper contour set and the lower contour set of (e_B, r_B) in terms of $E[y|e,r]$ is non-empty. In the upper contour set, revenue rises, while in the lower contour set, it falls.

To see whether the target T_M is greater or smaller than the baseline target T_B, note that the partial of (6) with respect to T is:

$$\begin{aligned} M(e,r,T) = 1 &- g(T|e,r)\left(1 + \frac{V}{U(e,r,T)}\right)\{T - \mathbb{E}[y|e,r,y \leq T]\} \\ &+ G(T|e,r)\frac{V}{U(e,r,T)^2}\frac{\partial U}{\partial T}\{T - \mathbb{E}[y|e,r,y \leq T]\} \\ &- G(T|e,r)\left(1 + \frac{V}{U(e,r,T)}\right)\left\{1 - \frac{\partial \mathbb{E}[y|e,r,y \leq T]}{\partial T}\right\} \end{aligned}$$

If $T \to \bar{y}$, then $M(e,r,T) < 0$, and if $T \to 0$, then $M(e,r,T) > 0$. By the intermediate value theorem, it is zero at some value in between. None of our assumptions restrict $M(e_B, r_B, T)$ to be positive or negative; hence it is ambiguous in general.

However, in case revenue falls $E[y|e_M, r_M] \leq E[y|e_B, r_B]$, the target has to fall as well. To see this, we apply the Implicit Function Theorem to $M(e,r,T) = 0$ at all (e,r) in the intersection of the lower contour set running through (e_B, r_B) and the lower quadrant given by $[e_b, \bar{e}] \times [0, r_b]$, which we denote by $\mathcal{S}_Q$. We then show that the total effect on the target of moving from (e_B, r_B)

to (e_M, r_M) is negative:

$$\begin{bmatrix} \frac{\partial T}{\partial e} \\ \frac{\partial T}{\partial r} \end{bmatrix} \cdot \begin{bmatrix} e_M - e_B \\ r_M - r_B \end{bmatrix} = - \begin{bmatrix} \frac{\partial M(e,r,T)/\partial e}{\partial M(e,r,T)/\partial T} \\ \frac{\partial M(e,r,T)/\partial r}{\partial M(e,r,T)/\partial T} \end{bmatrix} \cdot \begin{bmatrix} e_M - e_B \\ r_M - r_B \end{bmatrix} < 0$$

evaluated at all $(e, r) \in \mathcal{S}_Q$. This depends specifically on the the sign of the partials of $M(e, r, T)$. First, consider that lower expected revenue implies a lower optimal target. To see this, consider the lower bound: if expected revenue were near zero, then so is the target. Since, by assumption, revenue falls, the partial terms involving changes in revenue caused by the change in effort and risk are negative. Therefore, while the partial effect of the change in effort is negative and partial effect of the change in risk is positive, the sign of the dot product depends only terms that move with $U(e, r, T)$. Because we are moving away from the incentive compatible set, the effort-risk bundle is increasingly less favorable to the driver, lowering his valuation of the contract.

Finally, to see that the welfare effect is ambiguous, note that welfare is just profit minus disutility of work $\psi(e, r)$. While profit rises unambiguously, we do not know whether $\psi(e_M, r_M)$ is greater or smaller than $\psi(e_B, r_B)$ under the maintained assumptions, and in particular whether it overcompensates for the rise in profit.

□

B Data Collection

B.1 Sample Recruitment

We conducted an extensive recruitment drive in late 2015 by contacting owners through SACCOs that were operating along various routes across the city. All owners were informed at the time of recruitment that a monitoring device would be placed in their vehicle free of charge, and they would be required to *provide* daily information about their business operations. We also mentioned that a random subset of owners would be selected to *receive* information from the tracker via a smartphone app for a six month time period, while others would have to wait 6 months before gaining access to the information for a shorter two month period. Owners who expressed interest in the study during the recruitment drive were contacted again over the phone to confirm their willingness to participate in the experiment, and to check that they met the three study requirements (owners had to own a single matatu, which they rented to a driver, and manage the firm's operations themselves). In sum, it took approximately four months to recruit enough participants across 9 major commuter routes (Figure A.1b).

B.2 Installations

The first installation took place in November 2016, and continued until April 2017. The field team, managed by EchoMobile, was able to fit approximately 15 matatus per week with a device (Figure A.4). Each owner was compensated for the time their vehicle spent off-road to perform the installation of the device with a one-time payment of $50 (5000 KES) and a new Android phone (worth approximately $80) to ensure they could access the SmartMatatu app. The installation process was rather complex, requiring a team of three staff (an enumerator, a field manager, and an engineer). While the mechanic worked on fitting the device in the matatu, the field manager took the owner aside to re-explain the purpose of the research project and the tracking device's functionality. For owners in the treatment group, the field manager conducted an additional training on the app. At the same time, the enumerator administered the baseline survey to the driver in a separate location, outside of the owner's earshot, so that the driver felt comfortable answering the questions honestly. Once the field manager finished the training with the owner, and the enumerator finished administering the survey to the driver, they switched. The field manager then took the driver aside to explain that they would receive a daily SMS to elicit information about the day's operations and to emphasize that all of the data they shared would remain confidential. Meanwhile, the enumerator conducted a 20-minute baseline survey with the owner. This whole installation process took approximately 1 hour to complete. The field manager shared his contact information with the owner and the driver so they could contact him with any further questions they had.

B.3 Treatment Assignment

The first treatment arm is referred to as the "information treatment". Owners in our sample were randomly allocated to a treatment and a control group. Owners in the treatment group were provided with free access to the data produced by the monitoring device immediately after installation. Owners in the control group were informed that they would receive the same access six months after the device was installed. During the device installations our field manager spent an additional 30 minutes with treatment owners explaining the types of data that would be visible on the SmartMatatu app. A small survey was administered to the owners at the end of their training to make sure they knew how to find all the information contained in the app. Despite this in-depth training, it took owners a few months to feel comfortable navigating the different tabs in the app. We informed treatment and control drivers that a tracker would be placed in their vehicle. We did not mention, however, whether the information would be transferred to the owner. This meant that any subsequent changes we observed in driver behavior could only come from owners using

the tracker data, rather than from receiving different information from the enumerators during the installation.

Four months after the information treatment was launched, we introduced a second treatment arm, referred to as the "safety" treatment. We selected half of the treatment drivers and half of the control drivers and offered them cash incentives to drive safely. This arm was designed to simulate the role of a functioning regulatory system and monetize the tradeoff between revenue and safety that drivers face. The cash incentive drivers were then randomly split into two groups: a one-month treatment group and a two-month treatment group. This was done so we could study whether any changes in driving behavior that might be induced by the incentives would persist after they were removed. The specific incentive amount they received was determined by a safety rating, calculated daily for each driver in the following way. We analyzed two weeks of data for each driver (dropping days with less than 30km), tracking 1) the number of alerts of each type (speeding, heavy braking, sharp turning and over-acceleration), and 2) the number of hours worked. For each driver, day and alert type we computed the rate of violations by dividing the number of alerts by the number of hours worked. For each driver, we then constructed a distribution of these rates for each alert type and found the percentile into which that day's alert rate fell. We then calculated the weighted average percentile for each driver-day by adding the alert rates for each type, applying weights of 1/3 for speeding and braking, and 1/6 for over-acceleration and turning. The average we computed each day lay between 1 and 100. We assessed the cutoff below which they fell and disbursed their incentives accordingly (fewer safety violations resulted in a lower percentile and a higher payout).

We collected data from three different sources. The first data set is a panel of daily responses from owners and drivers which we gathered through the app and SMS surveys, respectively. Next the enumerators conducted 8 monthly surveys, beginning with the baseline, followed by 6 monthly surveys and wrapping up with the endline. Finally the GPS tracker collected a wealth of data that we use to measure driving behavior, including safety violations.

B.4 Non-system Application Variables

The SmartMatatu app was also designed to collect information from owners. Collecting accurate data can be very challenging in these settings, and this system was created to improve the quality of the data we received. Owners in the study were reminded daily via a notification on their phone to report on that day's business activities through a form located on the app. They were asked to submit data on: the "target" amount assigned to their driver at the beginning of the day; the amount the driver delivered to the owner; any repair costs incurred; an overall satisfaction score for their driver's performance (bad, neutral, good); and whether the driver was fired/quit that day. Once the report was successfully submitted, owners received $0.40 (40 KES) via M-Pesa (a mobile

phone-based money transfer service). We collected similar information from drivers through SMS surveys (because the drivers were not provided with smartphones). At the end of every work day around 10pm, drivers would receive a text message asking whether they were ready to respond to the survey. Once they agreed, individual text messages were sent to the driver asking for: the total revenue the matatu collected from fares that day; the amount they spent on fuel; and their "take home salary" (their residual income after expenses and paying the owner). Once the driver responded to all the questions, they were sent $0.40 via M-Pesa to incentivize consistent reporting.

We developed a set of processes for checking and validating the daily data we received from owners and drivers. Echo Mobile wrote code to check for anomalies including outliers and entries that did not make sense and/or suggested the owner/driver may not have understood how to answer the question. A team of enumerators would then follow up with owners and drivers over the phone about each one of these entries. In cases where owners and drivers were able to justify their responses, the enumerators would record their justifications in an excel spreadsheet. In cases when owners and drivers revised their responses, this data was corrected on the server.

B.5 Baseline and Follow-up Surveys

We conducted eight rounds of surveys. We first administered the baseline survey during the tracker installation. The *owner* baseline survey collected detailed information regarding basic demographics, employment history, features of the matatu, and their relationship with the current driver. Similarly the *driver* baseline asked about driver demographics, experience as a driver, unemployment spells, and their relationship with the current owner. For both owners and drivers we measured cognitive ability through Raven's matrices. We also used games to gauge drivers' risk aversion and driver/owner propensity to trust one another. To measure risk we asked respondents whether they would prefer to receive $5 (500 KES) for certain or play a lottery to win $15 (1500 KES). The game was repeated multiple times, with increasingly favorable lottery odds. The trust game presented owners with $5 (500 KES) and asked them to select a certain amount to be placed back in an envelope. They were informed that this amount would be tripled and delivered to the matatu driver who was then going to decide how much to keep for himself and how much to return to the owner. The amount they chose to place in the envelope was recorded in the survey. When playing the game with drivers, we first presented them with an envelope containing $9 (900 KES). This amount was standardized across all drivers to ensure they faced the same choice. The drivers were informed about the owner's decision and how this amount was then tripled. The drivers were asked to return however much they wanted to the owner.

We proceeded with 5 monthly follow-up surveys. The monthly surveys were administered with three purposes in mind. First, they provided an opportunity for enumerators to follow up regularly

with matatu owners and drivers and address any questions they might have about the device. Second, they were used to remind both parties to continue submitting the daily reports in the SmartMatatu app. Finally, they were designed to collect some basic data. As owners and drivers reached the 6-month mark, we conducted an endline survey to measure changes in key outcomes, and to assess the impact of the information treatment and the cash incentives.

B.6 Tracking Data

The CalAmp tracking device transmitted high frequency data on forward/backward/lateral/vertical acceleration, jerk, location and a timestamp. We use the raw measures of acceleration to investigate changes in driver behavior. Specifically, we look at vertical and lateral acceleration to determine whether the driver is operating on bumpier stretches of road. Furthermore, we use the GPS data to calculate how far each vehicle is from the route they are licensed to be on at any point in time. This provides a measure of how far the driver is deviating from the actual route. Figure A.2 depicts the number of times vehicles licensed to route 126 pass through a particular lon-lan cell. The first panel clearly shows what the route should be, and the second panel overlays the designated route to confirm. The figure illustrates that off-route driving is relatively common practice.

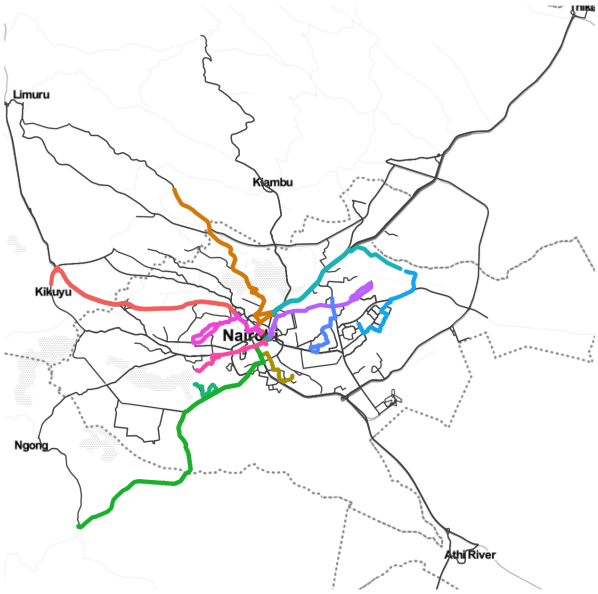
The tracker subsequently fed the raw data into an algorithm that computed the number of safety events that occurred in a 30 second time frame. Thresholds were calibrated for the Kenyan roads to avoid capturing an unreasonable number of safety violations and losing credibility among owners. These events included instances of speeding, over-acceleration, sharp braking, and sharp turns. The data was then further aggregated on the backend to produce daily reports on the number of safety violations, which is what we use for our analysis.

C Appendix Figures

Figure A.1: Metropolitan Nairobi matatu route maps

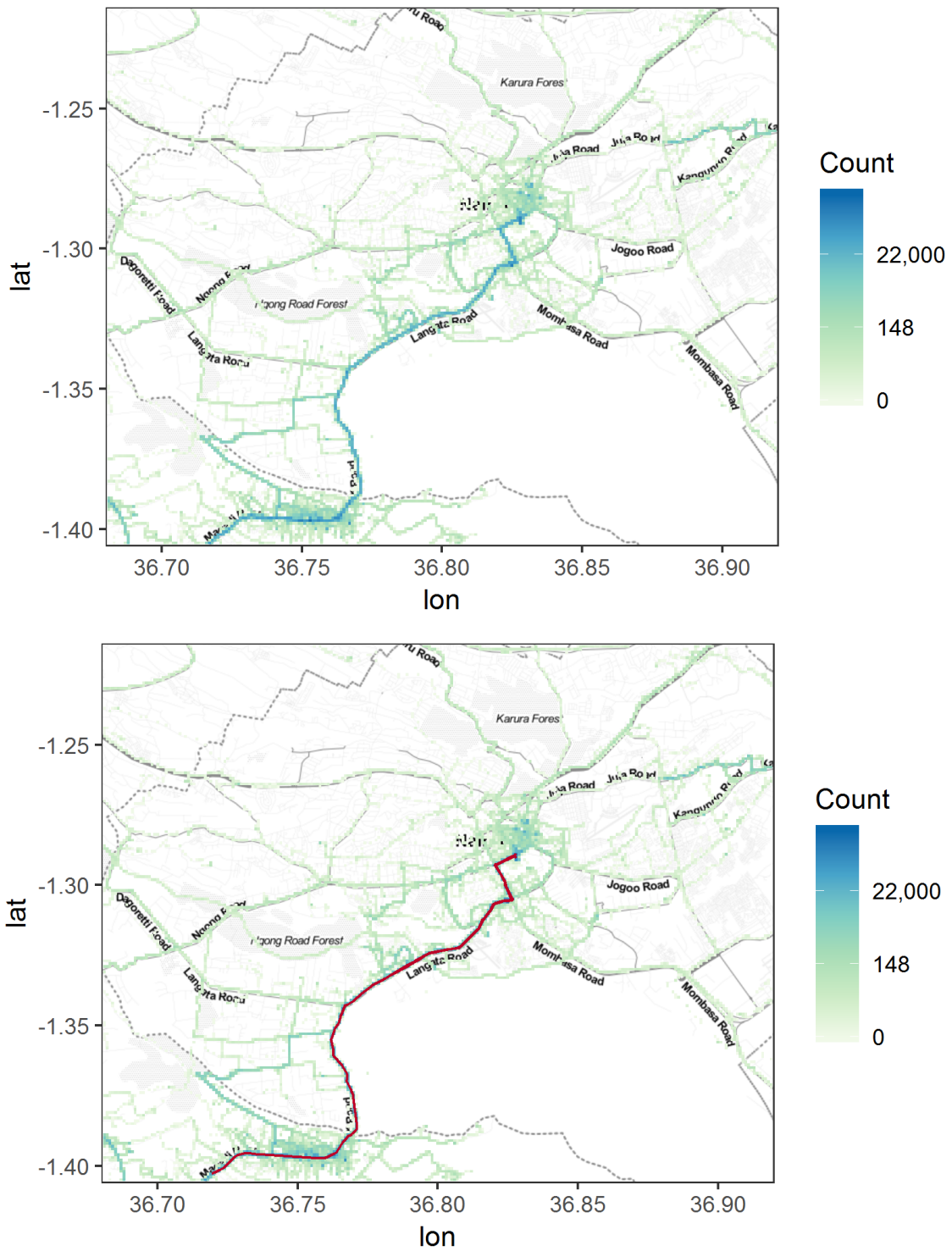
(a) Designated bus routes in Nairobi (black)

(b) Designated bus routes in Nairobi (black) and routes in our sample (colored)



Notes: Map of all routes in the Nairobi metropolitan areas. Panel A: all routes documented in the Digital Matatus project (`digitalmatatus.com`). Panel B: Our 255 participants are spread across the highlighted eleven routes.

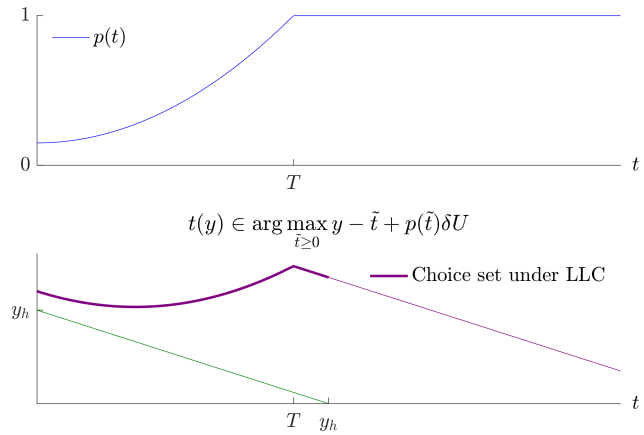
Figure A.2: Device location



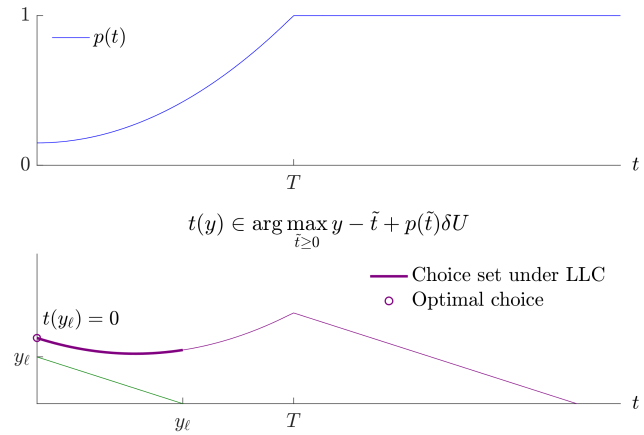
Notes: These maps use data from the trackers that were installed in vehicles licensed to operate on Route 126 (Ongata-Rongai line). We count the number of times that vehicles passed through particular longitudinal and latitudinal cells on the map. A deeper shade of blue demonstrates that more vehicles passed through that particular cell. The second panel overlays the designated route that vehicles are supposed to be on (red). Any colored cells outside of the designated route are instances of off-route driving.

Figure A.3: Setting the transfer and re-hiring schedules

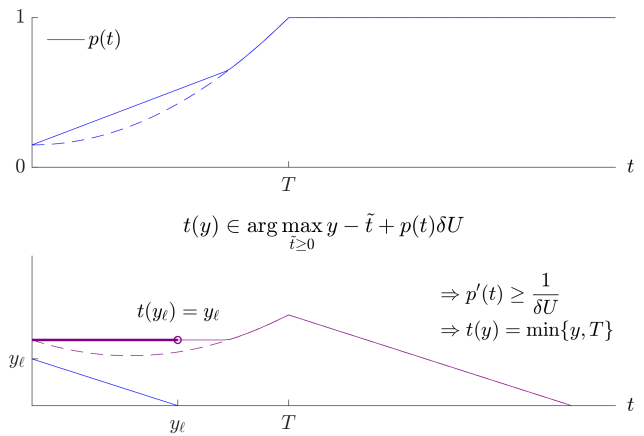
(a) An arbitrary rehiring schedule $p(t)$ and transfer under y_h



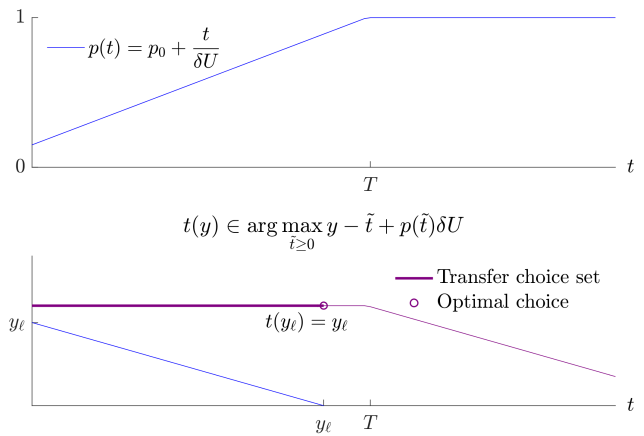
(b) Arbitrary $p(t)$ and transfer under $y_\ell < y_h$



(c) Improving arbitrary rehiring schedule.

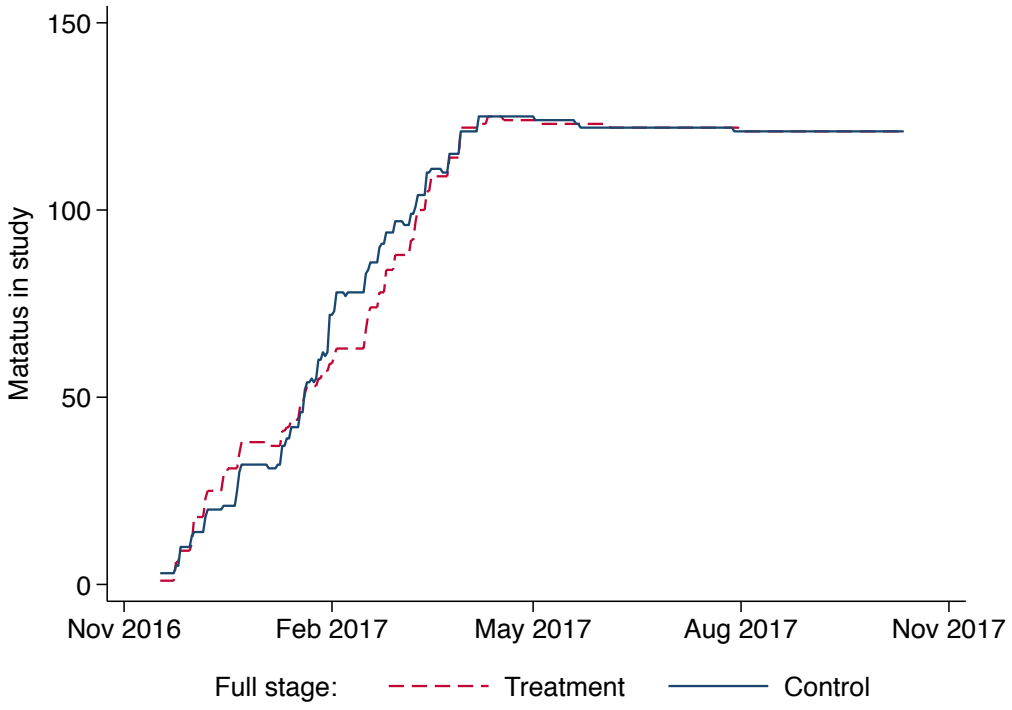


(d) Optimal rehiring schedule $p(t) = p_0 + \frac{t}{\delta U}$.



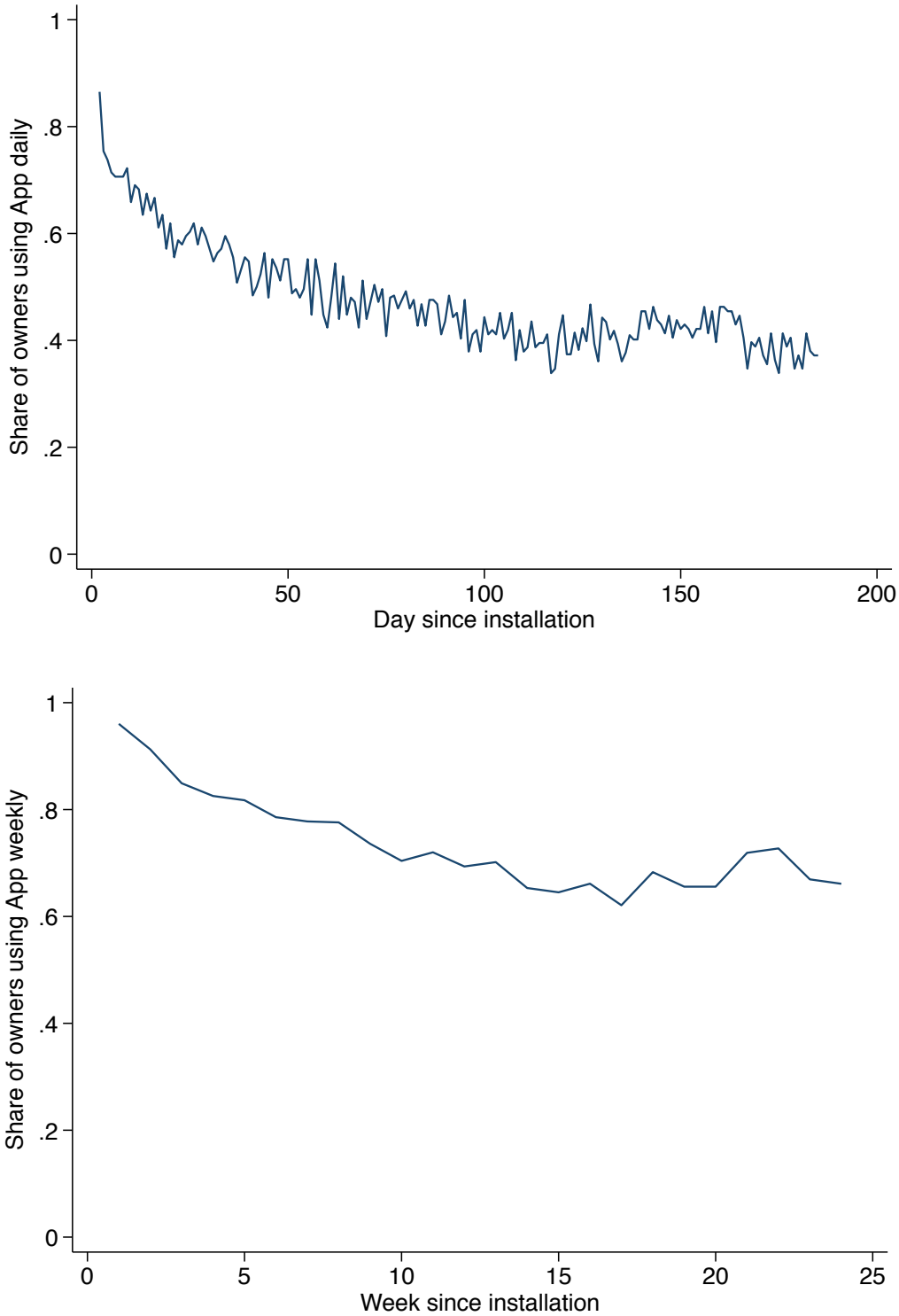
Notes: Intuition for the transfer and rehiring schedules in the Lemma. Panel A shows some arbitrary rehiring schedule with target T under a high realized revenue y_h. Under this rehiring schedule, the driver chooses the transfer that maximizes his utility in the transfer problem from the bold purple line; hence, he will choose to transfer T. If he instead only makes $y_\ell < y_h$, he will transfer zero under this rehiring schedule, as in Panel B. If the owner had instead guaranteed a minimal slope of $1/\delta U$ in the flat part of $p(t)$, the owner would have received y_ℓ instead, as in Panel C. A steeper slope offers no transfer benefits, and thus the owner would choose the schedule $p_0 + \frac{t}{\delta U}$ to maximize the driver's transfer to her.

Figure A.4: Installation timeline



Notes: Number of matatus that were fitted with tracking devices (and hence were added to the study) per week. The first installation took place in November 2016, and continued until April 2017. On average, the field team was able to fit trackers to 15 matatus per week. As a result it took approximately five months to finish installations.

Figure A.5: App usage since installation

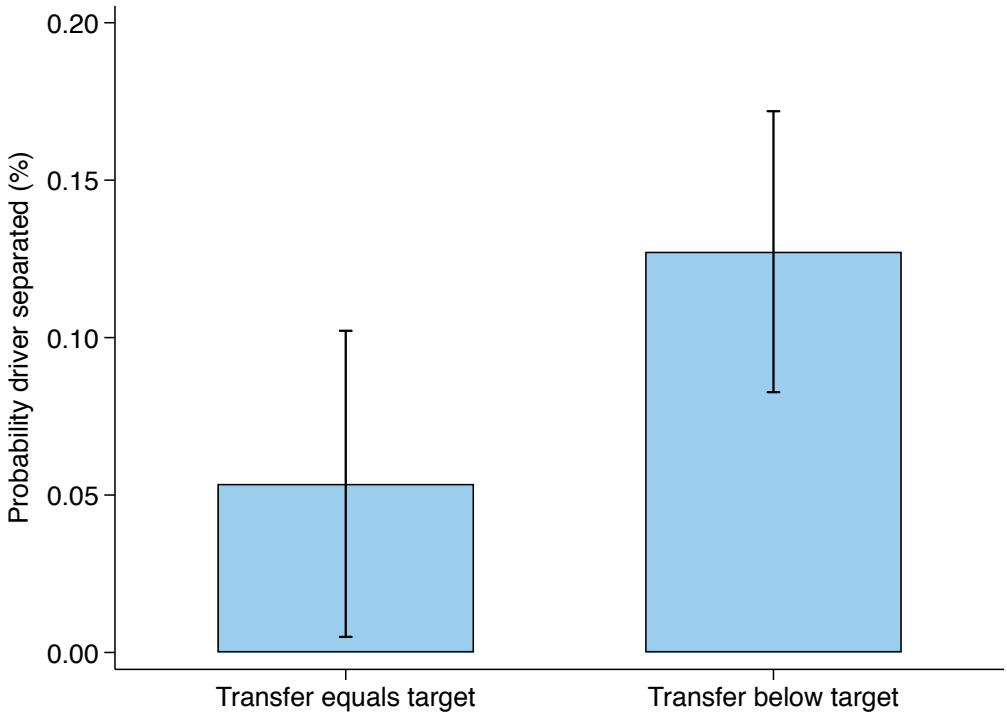


Notes: To measure device usage, we capture whether any API calls were made in a day. An API call is generated each time the owner requests data from the server, such as when logging in or refreshing a screen. The top panel looks at usage by week, whereas the bottom panel looks at usage per day.

Figure A.6: Example of vehicles used in study: 14-seater minibus

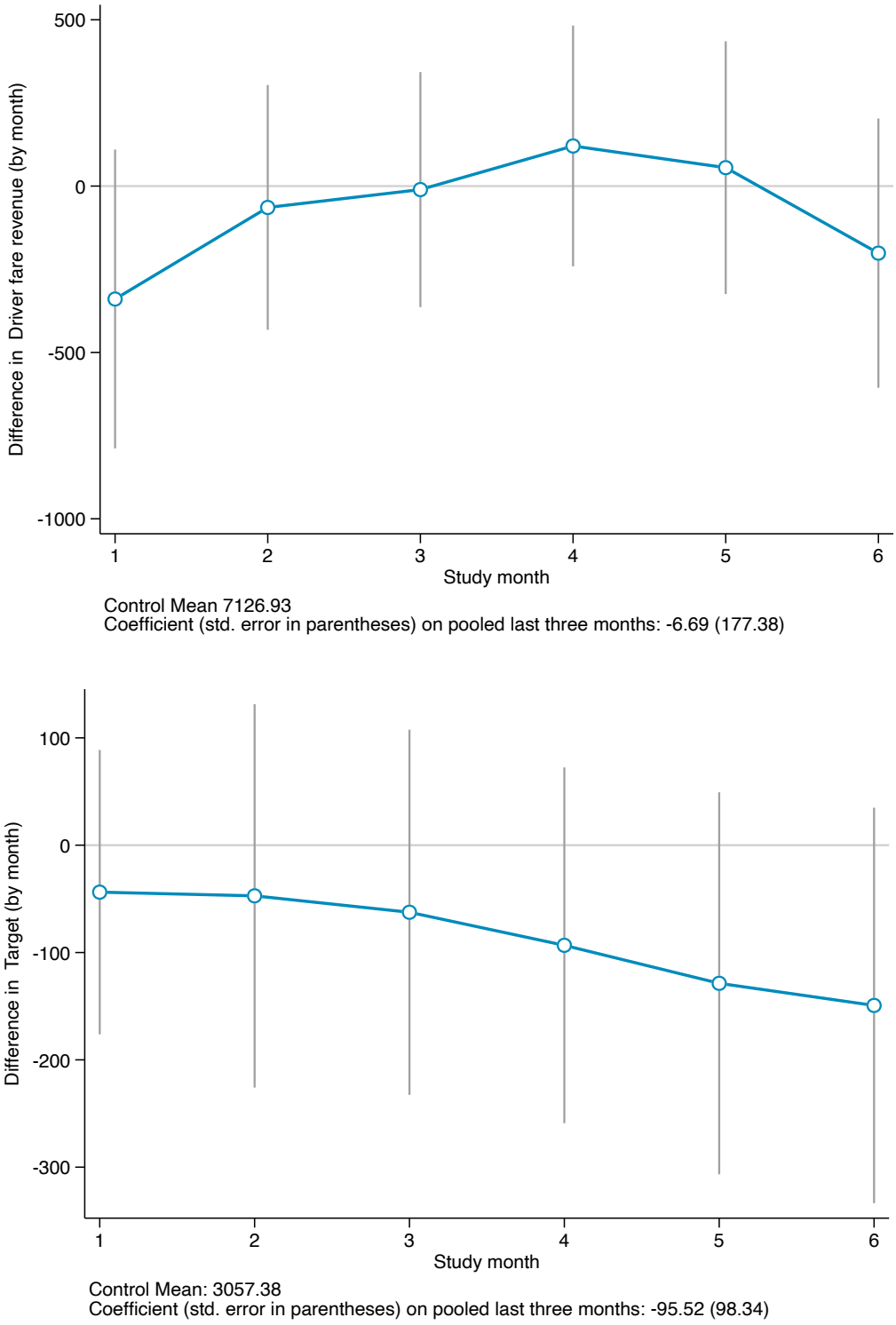
Notes: A typical 14-seater matatu in downtown Nairobi.

Figure A.7: Probability fired as a function of transfer



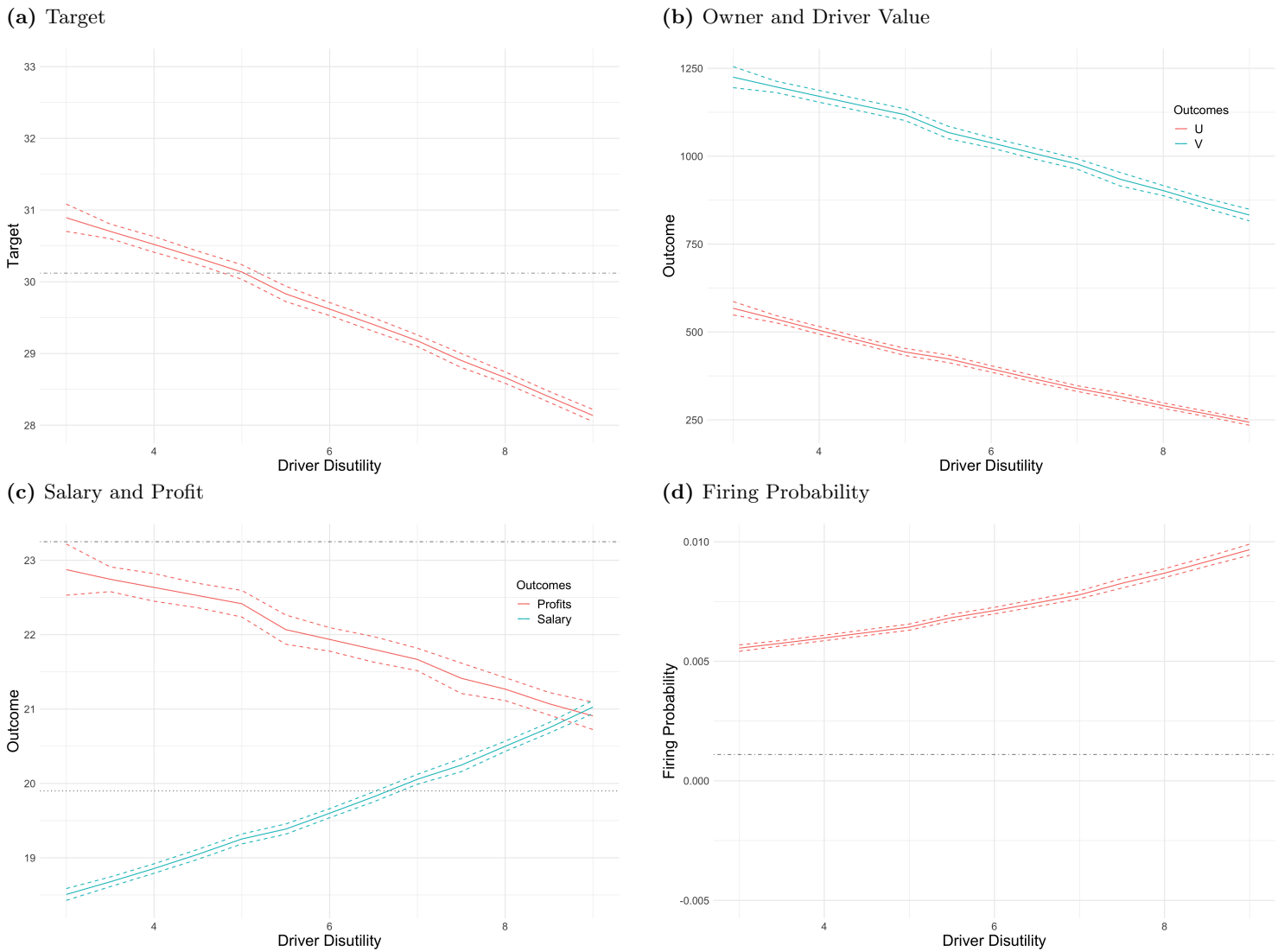
Notes: Daily share of drivers separated from owners, conditional on whether the transfer equals the target or is below the target. In total, 26 out of 255 drivers separated from their owner over the six month period. xxx add this code to analyze_data.do

Figure A.8: Effect of monitoring on revenue and target



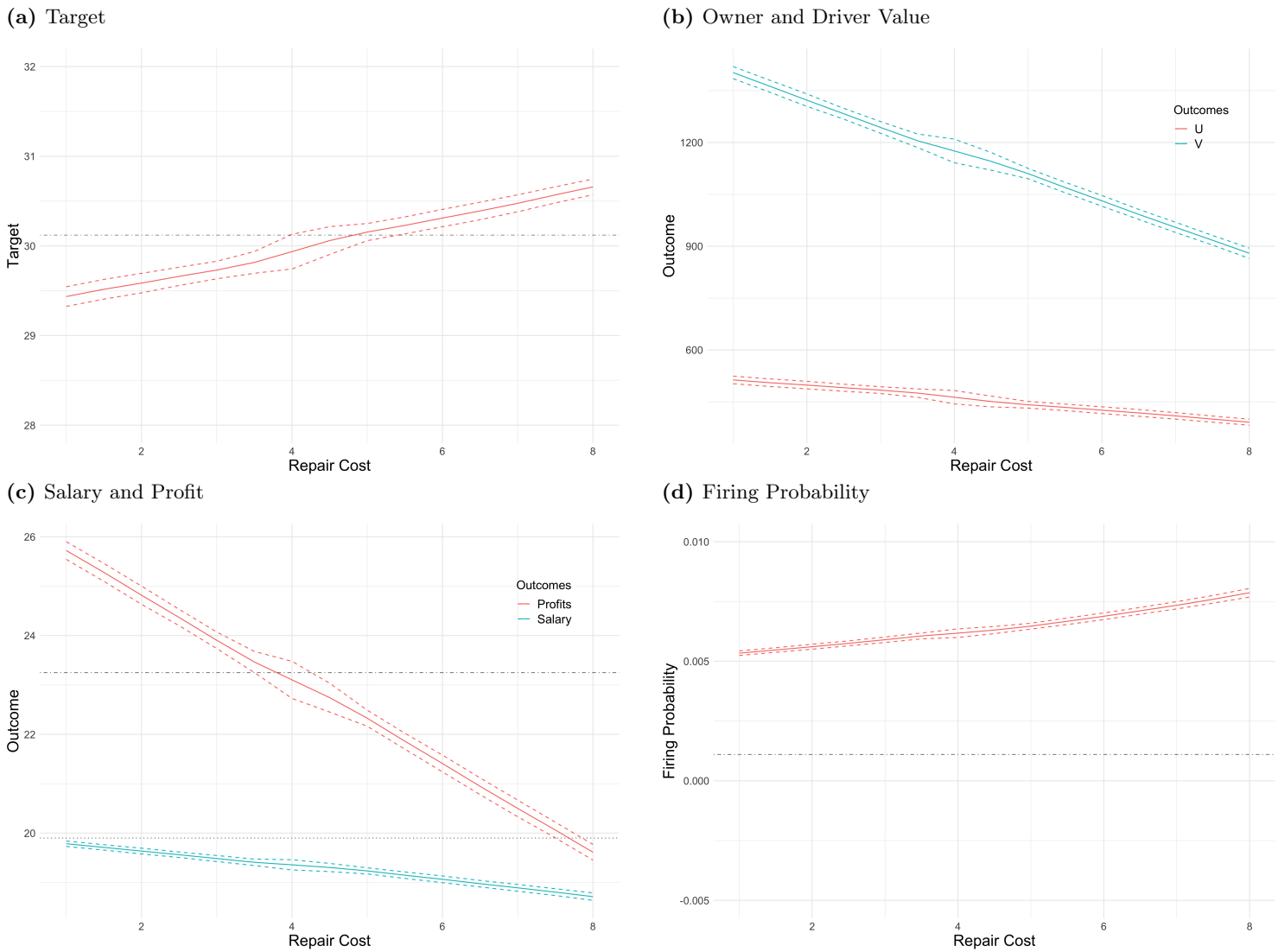
Notes: OLS estimates according to Equation 2. Treatment effects by month on daily revenue. Standard errors for 95% confidence intervals clustered at the matatu level (255 clusters).

Figure A.9: Model Calibration Sensitivity to Work Costs



Notes: Figures plot the model predicted values for different input values of driver work cost ($\psi(e, r)$). Dashed or dotted lines show observed values where applicable. Other input parameters are fixed per those stated in the main model calibration table.

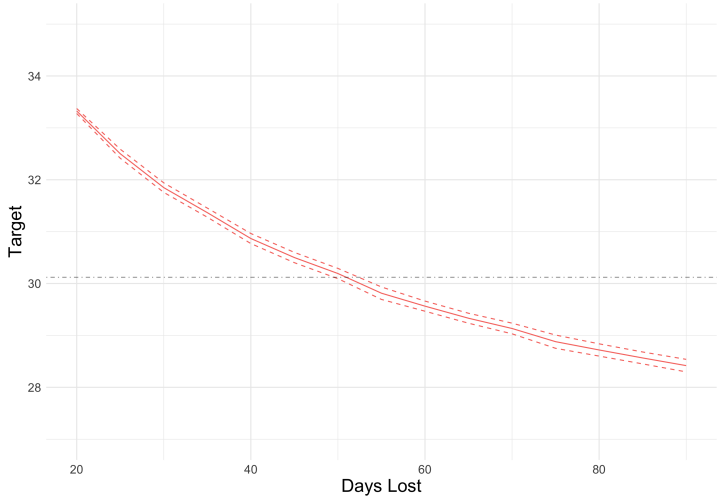
Figure A.10: Model Calibration Sensitivity to Repair Costs



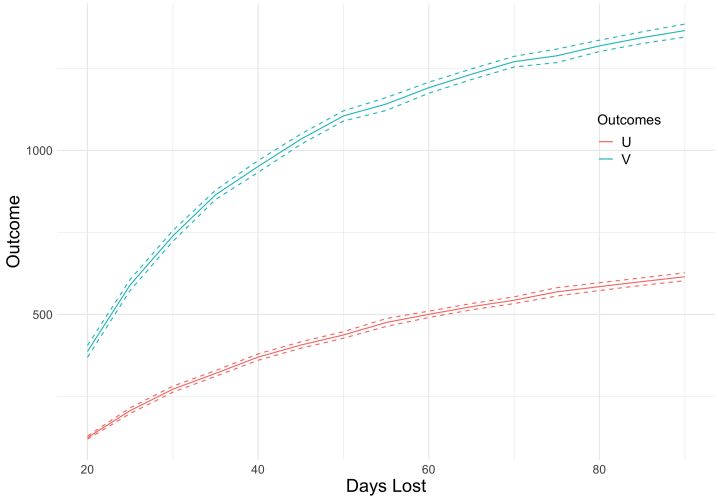
Notes: Figures plot the model predicted values for different input values of expected repair costs ($\mathbb{E}[c(r)]$). Dashed or dotted lines show observed values where applicable. Other input parameters are fixed per those stated in the main model calibration table

Figure A.11: Model Calibration Sensitivity to Firing Costs

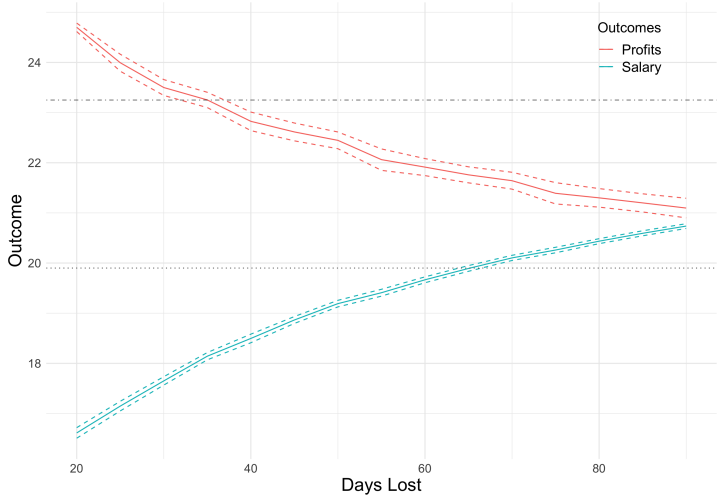
(a) Target



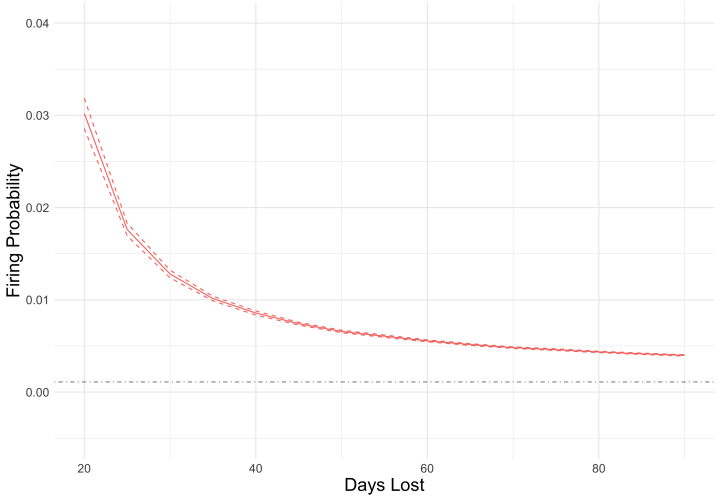
(b) Owner and Driver Value



(c) Salary and Profit

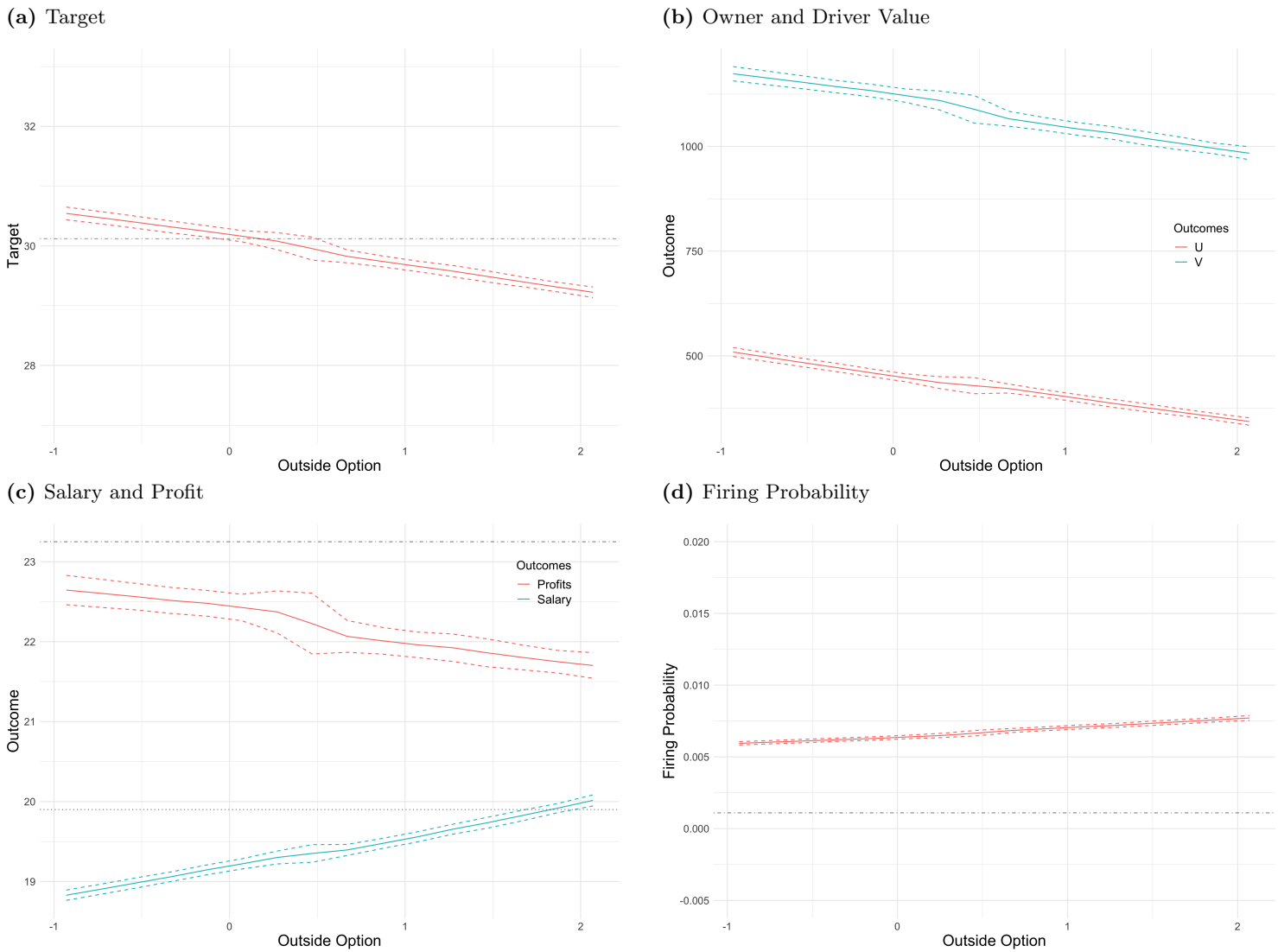


(d) Firing Probability



Notes: Figures plot the model predicted values for different input values of firing costs (h). Dashed or dotted lines show observed values where applicable. Other input parameters are fixed per those stated in the main model calibration table

Figure A.12: Model Calibration Sensitivity to Outside Option



Notes: Figures plot the model predicted values for different input values of outside option (u). Dashed or dotted lines show observed values where applicable. Other input parameters are fixed per those stated in the main model calibration table

D Appendix Tables

Table A.1: Treatment effects on target, costs, profits, mileage and hours worked

	(1) Target	(2) Repair costs	(3) Gross profit	(4) Mileage (kilometers)	(5) Device on (hours)
Treatment × Month 1	-43.74	67.97	-181.60	-4.28	-0.84
	(67.59)	(69.96)	(238.35)	(5.96)	(0.66)
Treatment × Month 2	-47.19	-50.18	63.27	2.13	0.63
	(91.13)	(73.55)	(208.06)	(5.60)	(0.59)
Treatment × Month 3	-62.48	-124.21	89.19	7.94	1.00*
	(86.78)	(79.55)	(218.80)	(5.31)	(0.55)
Treatment × Month 4	-93.27	-184.97**	449.04**	4.56	0.71
	(84.56)	(89.35)	(223.61)	(5.69)	(0.64)
Treatment × Month 5	-128.68	-180.89*	453.49**	9.51	1.45**
	(90.77)	(93.37)	(213.32)	(6.41)	(0.72)
Treatment × Month 6	-149.30	-215.73**	179.75	12.93*	1.45*
	(93.99)	(102.80)	(227.27)	(6.90)	(0.76)
Control Mean of DV	3057.38	483.48	3260.50	96.64	14.79
Controls	X	X	X	X	X
Day FE	X	X	X	X	X
Route FE	X	X	X	X	X
Matatu N	237	238	216	254	254
Matatu-Day N	15,888	15,881	10,406	45,654	45,654

Notes: OLS regressions as in Equation 2. "Target": daily revenue target set by owner. "Repair costs": owner-reported daily repair costs. "Gross profit": Revenue minus repair costs minus driver residual claim (salary). "Mileage (kilometers)": Daily mileage as measured with tracking device. "Device on (hours)": number of hours the tracking device reported the ignition to be on as a measure of driver work hours. Controls include the age of the matatu, the number of special features, owner age and sex, owner education, owner self-employment status, the number of other businesses the owner runs, owner years of matatu industry experience, and owner raven score. Data are from daily panel collected from owner in-app reports and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level. Stars indicate statistical significance at the 1% ***, 5% **, and 10% * levels. xxx express all values in dollars instead SEK.

Table A.2: Perceptions of treatment at endline

	(1) Trust amount	(2) More honest	(3) Performance rating
Treatment	33.80**	0.71***	0.11
	(15.12)	(0.05)	(0.17)
Control Mean of DV	151.61	0.04	7.21
Controls	X	X	X
Route FE	X	X	X
Matatu N	244	190	246

Notes: OLS regressions of outcome on treatment indicator, controlling for route fixed effects, the age of the matatu, the number of special features, owner age and sex, owner education, owner self-employment status, the number of other businesses the owner runs, owner years of matatu industry experience, and owner raven score. "Trust amount": amount in KES the owner gives to the driver in a trust game at endline. "More honest": change in honesty of transfer amount since baseline, either less honest (-1), the same (0), or more honest (1). "Performance rating": overall performance rating of the driver at endline, ranging from 1 (poor) to 10 (excellent). Data from endline survey; trust and performance rating added after a quarter of endlines were already completed, hence only up to 190 out of 255 observations (balanced across treatment and control; three non-responses to first three questions). Robust standard errors. Stars indicate statistical significance at the 1% ***, 5% **, and 10% * levels..

Table A.3: Simulated vs. Actual Treatment Effects

Panel A: Assumptions

Input	Symbol	Value	Justification
Repair Costs	$E[c(r)]$	2.7	Previous value of 4.9 minus observed treatment effect reduction of 2.2
Revenue Distribution	$G(\cdot)$	—	Empirical distribution calculated from data
Discount Factor	δ	0.99	
Firing Cost	h	51.44	Fixed from previous value
Outside Option	$\bar{u}$	7.07	Fixed from previous value

Panel B: SMM Parameter Estimates

Input	Symbol	Value	Interpretation
Disutility	$\psi(e^*, r^*)$	4.53	Decrease of 0.46 (9%) from previous value
		(0.03)	

Panel C: Predicted vs. Observed Treatment Effects

Outcome	Pred.Treat	Obs.Treat	Difference
Target	-0.47	-1.13	0.66
	(0.07)	(0.87)	(0.87)
Expected Profit	1.6	3.62	-2.02
	(0.13)	(1.96)	(1.96)
Expected Salary	0.27	0.23	0.04
	(0.05)	(0.35)	(0.35)
Prob. Separation	-0.0005	-0.01	0.01
	(0.00008)	(0.01)	(0.01)
Driver Contract Value minus Outside Option	63.06	—	—
	(7.46)		
Owner Value	130.98	—	—
	(11.65)		
Total Welfare	194.04	—	—
	(9.84)		

Notes: All values are in 100s of Kenyan Schillings. Work cost estimated via simulated method of moments matching the observed target, salary, and profit treatment effects.