Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit

Erin M. Kelley Gregory Lane David Schönholzer*

October 20, 2023

Abstract

Small firms struggle to grow beyond a few employees. We introduce monitoring devices into commuter minibuses in Kenya and randomize which minibus owners have access to the data using a novel mobile app. We find that treated vehicle owners modify the terms of the contract to induce higher effort and lower risk-taking from their drivers. This reduces firm costs, and increases firm profitability. There is suggestive evidence that some firms expand. These results suggest that small firms may be able utilize monitoring technologies to overcome problems of moral hazard and enhance their profitability.

JEL codes: O12, O18, D86.

*Kelley: DIME, World Bank (erinmkelley@worldbank.org). Lane: Harris School of Public Policy, University of Chicago (laneg@uchicago.edu). Schönholzer: IIES, 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, David McKenzie, Edward Miguel, Xiaosheng Mu, Matt Pecenco, Marco Gonzalez-Navarro, Kate Pennington, Louis Preonas, Manaswini Rao, Mark Rosenzweig, Elisabeth Sadoulet, Larry Samuelson, Jakob Svensson, Ralph Winter, Chris Woodruff and Matt Woerman for helpful comments and suggestions. This project received IRB approval from UC Berkeley (#2015-02-7180) and the University of Nairobi (#KNH-ERC/A/417). We gratefully acknowledge financial support from DIL and PEDL (#3022, #3774, #3839). We are also extremely grateful to the entire team at EchoMobile (Jeremy Gordon, Boris Maguire, James Mwangi, Vitalis Omutanyi, James Langat, James "Wambururu" Kariuki, and Kevin Onyango) 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.

1 Introduction

Small and medium sized firms account for the majority of businesses in low-income countries, they employ over half of the population, and account for more than 40% of GDP (World Bank, 2021). Firms in low income countries also appear to stay small suggesting they face barriers to growth (Hsieh and Olken, 2014). Yet, the transition from an economy dominated by many small firms operated by small-scale entrepreneurs to one with larger firms that promote wage employment is an important step in a country's economic development (McKenzie, 2017; Gollin, 2002). Therefore, identifying and relieving the many constraints that firms face has become a central focus for researchers and policymakers alike (Bloom et al., 2013; McKenzie, 2017; Brooks et al., 2018; Hardy and McCasland, 2020).[1]

Any firm seeking to expand needs to grapple with the challenges associated with managing their workforce. Seminal work by Lucas (1978) argues that firm size depends on the number of workers that managers can effectively supervise (their "span of control"). When firms cannot observe all dimensions of their employees' behavior, problems of moral hazard emerge, which may harm firm productivity and lower profits. If firms do not have systems in place to effectively monitor their workers, their span of control may be limited, and their ability to scale their business may be reduced. Theoretical work has explored this mechanism: Shahe Emran et al. (2021) and Akcigit et al. (2021) suggest that small firms do not expand because hiring, delegating responsibilities, and supervising someone else is much more costly than the owner doing the work themselves. While empirical work suggests that problems of moral-hazard and lack of trust within the firm abound, especially in low-income countries (Bassi et al., 2022; Caria and Falco, 2022), there is little evidence on whether these challenges impact firm productivity, and whether potential solutions are effective (Jayachandran, 2020).

In this paper we investigate the causal impact of one such solution on worker behavior and firm outcomes. Specifically, we document whether the use of monitoring devices helps *small* firms change their contract, boost profits, and stimulate growth. We do so among a population of firms with one employee in Nairobi. We focus on Kenya's informal public transit industry, which is dominated by privately run minibuses. These private firms struggle to grow beyond one or two employees, and problems of moral hazard exist. Minibus owners cannot observe how much revenue the minibus driver collects in passenger fares or whether he drives recklessly, which puts passengers at risk and increases vehicle repair costs. If owners are unable to observe the behavior of their employees, making it difficult to attribute poor

[1] See this IGC report for a comprehensive review.

business outcomes to their actions, problems of moral hazard can limit firm performance and persist for some time. This is especially true if the pool of alternative employees that owners can choose from is limited. In a descriptive survey we conduct among minibus owners, 75% of owners say they sometimes or always face challenges monitoring their workers. Relatedly, 70% of the owners agree that less than half of the matatu drivers in the industry are good or reliable, which implies that replacing a driver is a difficult and risky endeavor.

In this context, monitoring technologies can enhance outcomes by granting owners increased visibility into their employees' operations.[2] To understand how monitoring affects these firms, we develop a new monitoring system tailored to the industry that tracks driver effort and risk-taking choices. Our technology reveals driver actions but not revenue directly because it does not capture how many passengers are in the vehicle. Specifically, the system reports the driver's location, hours worked, distance driven, and a number of safety violations. We fit 255 minibuses with these tracking devices, working exclusively with owners who only manage a single minibus.[3]. 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. Drivers in both groups are told that a tracking device is fitted in their vehicle but it is up to treated owners to reveal that they have access to the information from the device.[4]

Interpreting the impacts of monitoring technologies requires observing and understanding the relationship between firms and their employees. In many informal transit systems around the world, minibus owners hire 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, but he may not be rehired for the next day if he transfers less than the target; and the owner is liable for major expenses accrued during the day. This *target contract* is not unique to Kenya, and can be found between minibus owners and their drivers across the world (Cervero and Golub, 2007; Bruun and Behrens, 2014).

We develop a model that shows how this target contract is optimal given the constraints owners face, but is inefficient from a social planner's perspective.[5] In other words, we show

[2]It is important to note that because these firms regularly engage with a single employee the impact of monitoring technologies might be more limited than in other contexts. Nevertheless, we anticipate that monitoring technologies will continue to impact firms' operations as long as there remains uncertainty about which poor business outcomes can be directly attributed to employees' performance.

[3]Approximately one third of owners in our baseline sample of the industry in Nairobi owned one minibus.

[4]Most drivers were familiar with tracking devices because of their use for security reasons.

[5]The principal-agent model we develop accounts for the myriad of contracting constraints frequently encountered in settings with relational contracts. First, the owner cannot observe effort and risk choices. Second, the owner cannot observe the amount of revenue the driver collects, and can only rely on the

that alternative contract arrangements such as debt contracts or wage contracts are suboptimal in environments where output is unobserved and drivers face limited liability. While the target contract is optimal, it incentivizes excessive risk-taking *and* high effort. This is because the principal cannot contract lower risk-taking in ways that are incentive compatible for the driver and maintain the flow of transfers from the driver to the owner. Monitoring technologies expand the contract space by making 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, resulting in lower costs to the firm.

Our results are consistent with these predictions. We find that treated owners are able to effectively use the system to monitor their driver's activities more easily. Owners retain the target contract structure, but there is some indication that the parameters of this contract change (albeit imprecisely estimated). By the end of the experiment, owners have slightly lowered the driver's daily revenue target by 4.9% (p-value = 0.114), and driving behavior is geared towards more effort and less risk-taking. Treated drivers increase the number of hours they spend on the road by 9.8% (p-value = 0.055) but engage in substantially less costly behavior such as off-road driving (p-value = 0.022), earning about the same amount of revenue as before. This lowers repair costs by 44.6% (p-value = 0.037) and contributes to substantial increases in daily profit for the owners. After four months, profits increase by 13.7% (p-value = 0.046), which is primarily driven by lower repair costs.[6] These gains in firm profits more than offset the cost of the device, suggesting that a tracking device like the one we designed for this study would be a worthwhile investment if it were available on the market. Finally, we investigate whether firms use these technologies to expand their business. We find weak statistical evidence that treatment owners are 12.9 percentage points more likely to own an additional vehicle (a 10.5% increase) than control owners by the end of the study (p-value = 0.091). These results suggest that monitoring technologies have an impact even when owners and workers interact regularly and problems of moral hazard may be less severe.

As these technologies become increasingly widespread, various institutions have expressed their concerns over their distributional consequences (West, 2021). We explore this by estimating the welfare implications of these devices. To quantify owner and driver welfare under the status quo and with the introduction of monitoring, we estimate the structural parameters of the model via generalized method of moments (GMM) using data from our

transfer from the driver to determine whether to rehire him for the next day. Third, drivers are often liquidity constrained and thus subject to limited liability. Finally, contracts need to be self-enforcing.

[6] Profits remain higher in month five but fall in month six.

experiment. We first estimate driver and owner welfare with data from the control group. Our estimates suggest the present-discounted contract surplus is large: the driver values the contract at about \$507 and the owner at about \$2,177. We then apply our GMM procedure to estimate the welfare effects under monitoring. Matching on reduced form moments from the experiment, we estimate that the owner gains about \$83 from higher profits under similar revenue. This is similar to their average willingness to pay of \$45 for the device at the end of the study. On the other hand, the driver's present-discounted value of the contract falls by 4% due to monitoring. This is primarily because they incur more disutility from having to drive in a less risky way. Thus, while the impact on welfare changes is imprecisely estimated, the total welfare effect of monitoring is small. These welfare estimates do not factor in the benefits of a better working relationship between owners and drivers: in a survey we conducted six months after the experiment, 98% of drivers said they preferred driving with the device because it improved the level of trust with the owner, and owners devote 30% more to drivers in a trust game at endline.[7]

Our paper demonstrates that monitoring technologies can help small firms align their employees' incentives with their own, reducing firm costs and boosting profits, with suggestive evidence that firms expand. Yet, it is important to note that we are studying the impact of monitoring technologies in a specific environment where owners have one worker they interact repeatedly with, whose behavior they cannot always easily observe. As such, our results are moderated by the size of firms we work with, and the pre-existing employee-employer relationship. First, the firms we work with only have one employee, and one valuable asset. To the extent that problems of moral hazard are larger among firms that have multiple workers and assets to supervise, the impacts we observe may be smaller than what we would expect to see elsewhere.[8] Second, the firms we work with have an ongoing relationship with their employees. To the extent that repeated interactions make it easier to attribute poor business outcomes to an employee's performance, and poor performers can be easily replaced, the scope for monitoring may be reduced.[9]

[7]We further discuss driver welfare considerations in Section 8. Besides changes in the working relationship, our welfare evaluation also abstracts from welfare effects on passengers – an important consideration in light of how dangerous minibuses are. We explore this in a companion paper (Lane et al., 2022). Monitoring did not significantly affect the number of accidents.

[8]On the other hand, if owners have more workers to supervise, and more assets to protect, they may have less bandwidth to utilize the information they receive from the device.

[9]On the other hand, while owners and drivers interact daily, owners are unable to be physically present beside their drivers. This could make it more difficult to directly attribute poor outcomes to employee performance relative to other settings such as factories, where managers can closely observe their employees' work.

Nevertheless, we see this paper as providing a proof of concept that moral hazard can impact firm profitability, and monitoring technologies may help some firms become more profitable, even among small firms who know their employees well. While the firms we work with are indeed small, they represent a common class of firms in Iow-income countries: 99 percent of the firms in many low-income countries have 10 workers or fewer (McKenzie and Paffhausen, 2019). Similarly, generalized levels of trust remain low among small businesses where employees and owners interact regularly. Caria and Falco (2022) show that generalized levels of trust are low in low-income countries, and as generalized levels of trust in society increases, the share of new businesses that expect to expand grows. Finally, monitoring capacity remain limited for many firms in low-income countries, implying there is scope for monitoring technologies to affect change.

Our study contributes to a number of literatures. Our work speaks to a large literature documenting barriers to firm growth in low-income 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. Most of the work in this field studies the impact of interventions that train firms on how to manage aspects of the business that do not involve employees (Bloom et al., 2013; Berge et al., 2015; McKenzie and Woodruff, 2017). Yet, a few papers provide evidence that managing employees is a challenge in low-income countries. Caria and Falco (2022) conduct a lab-in-the-field experiment to show that small businesses in Ghana do not trust their workers, which discourages them from hiring, while Bassi et al. (2022) study rental markets in Uganda and speculate that small firms engage in complex rental arrangements with one-another to avoid merging and dealing with the challenges of recruiting and managing a larger labor force. Atkin et al. (2017) show that firm and employees' mis-aligned incentives lead to sub-optimal technology adoption within the firm. Our paper focuses directly on employee management, documenting how the provision of information to the firm about employee behavior can affect informal contracts and the profitability of the firm.

Our paper also contributes to a growing literature studying the importance of monitoring. Anecdotally it has long been recognized that monitoring technologies are useful: larger firms in high-income countries are increasingly reliant on these tools: according to a study by the American Management Association, 80 percent of major companies monitor internet usage, phone and email of their employees (American Bar Association, 2018). By merging data from the World Management Survey to the World Bank's Data Catalog for GDP, we can also show correlational evidence that firms are more likely to use meaningful metrics to track employee performance as GDP improves (Appendix Figure A.1). Firms' substantial investment in monitoring tools in higher-income countries may suggest these technologies can play a role in facilitating firm growth.

There are a few papers that *empirically* estimate the impact of monitoring. We build on seminal empirical work by Hubbard (2000, 2003) and Baker and Hubbard (2004) who investigate how the introduction of onboard diagnostic computers affected the U.S trucking industry. Our study differs in three important ways from this earlier work. First, we generate exogenous variation in the usage of monitoring technologies by randomizing which companies receive data from a tracking device. Second, we 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. Finally, we study the impact of monitoring in a low-income country context, where relational contracts are more prevalent and monitoring devices reduce rather than eliminate information asymmetries. In this setting, we show empirically and theoretically that the impact of monitoring is fundamentally different. In closely related work, de Rochambeau (2020) finds that monitoring induces Liberian truck drivers to supply higher effort, although her focus is on intrinsic motivation.

More broadly, there exist a set of papers that document the importance of performance-based monitoring within the firm in high income countries. Gosnell et al. (2020) find that performance monitoring of Virgin Atlantic's airline captains improves labor productivity. Liu et al. (2021) also find that monitoring systems in Uber reduce driver moral hazard. Finally, Gertler et al. (2023) make individual performance of procurement agents available to managers and find significant reductions in overspending. Prior research suggests that monitoring systems may affect firms differently in low-income countries than in high-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 firms' ability to utilize the information they gather from new technologies. Our work demonstrates that this is not the case, but underscores how the impact of monitoring may be an important constrain to firm profits in low-income countries.

Finally, there is growing recognition that the informal transit industry has a large impact on various development outcomes. Research in this literature has focused on documenting the impact of policies that improve the efficiency of transportation networks within cities. This includes work by Hanna et al. (2017) and Kreindler (2020) that investigate the benefits of traffic congestion management policies. Tsivanidis (2019) documents large aggregate welfare gains from the introduction of the world's largest Bus Rapid Transit system in Colombia. Recent work by Lane et al. (2022) focuses on the impact of policies that improve the safety of informal transit systems. This builds on previous work by Habyarimana and Jack (2011, 2015) who study how to mobilize passengers to improve minibus safety.

The rest of the paper is organized as follows. In Section 2 we explain the rationale behind focusing on firms in Nairobi and provide evidence that the challenges they face are not unique to this context. In Sections 3 and 4 we describe the monitoring technology, the data collection, and the experimental design. Section 5 then develops a theory of contracting in this industry. We present reduced-form results of the experiment in Section 6. Section 7 provides results from our structural estimation, and Section 8 contextualizes its welfare implications. Finally, Section 9 concludes.

2 Context: Moral Hazard and Firms

In this section, we provide an overview of the matatu industry and explain why it is a relevant setting for studying the impact of moral hazard and monitoring technologies on firms. We also discuss how some features of this setting may generalize to other contexts.

2.1 Moral Hazard

Informal transit systems play a vital role in the public transportation of low-income countries, often accounting for more than two-thirds of daily commutes (Godard, 2006). In Kenya, these informal transit services, particularly the privately-owned minibuses known as *matatus*, are indispensable. Rough estimates suggest that approximately 15,000 matatus operate within the city, supporting a massive industry that employs over 500,000 people and contributes up to 5% of the country's GDP (Kenya Roads Board, 2007). Private entrepreneurs dominate this sector, typically owning a small fleet of vehicles. Passengers can board these matatus at various points along their route and pay their fare in cash (Bruun and Behrens, 2014). In Kenya, private entrepreneurs typically purchase 14-seat minibuses and obtain licenses for operating on specific routes.[10] The management of routes is overseen by different Savings and Credit Cooperatives (SACCOs). While the daily operations of the vehicles are handled by

[10]Matatu fares vary between $0.5 and $1.5 for travel inside the city center, and between $1 and $5 for trips to the outskirts.

their respective owners, the SACCOs play a crucial role in coordinating centralized activities. This includes addressing internal conflicts among owners and ensuring adherence to the regulations set by the National Transport and Safety Authority (NTSA).

Owners typically hire a single driver to operate their vehicle along the designated route.[11] 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. Drivers are hired on target contracts: the owner sets a daily revenue target at the beginning of the day and the driver is the residual claimant.[12] 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 can 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 the beginning of the school year) (Behrens et al., 2015).

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, 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. Similarly, the presence of severe competition within a route leads to reckless driving and high vehicle maintenance costs. According to the World Health Organization's Global Status Report on Road Safety, approximately 3,000-13,000 people die annually from traffic incidents in Kenya, and at least 62% of cases involve matatus (Odero et al., 2003; WHO, 2013).

To document the extent of moral hazard in this environment, we conducted focus group discussions prior to the experiment, and a series of descriptive surveys five years later in 2022 with 150 matatu owners operating on different routes across Nairobi. The sample for the 2022 descriptive survey was selected based on very general criteria: the owners had 1-5 matatus, and were willing to spend 30 minutes and receive $3 as compensation. We find that matatu owners are unable to observe their drivers' behavior with complete accuracy. According to our survey, approximately 65% of owners state that they can only sometimes

[11] Transit industries characterized by privately owned minibuses that establish target contracts exist in various other countries, including Mexico (peseros), the Philippines (jeepneys), Indonesia (tuk-tuks), India (rickshaws), and Tanzania (dala-dalas), among others.

[12] In Kenya, the driver is accompanied by a fare collector who the driver appoints and they work as a team (there is no cross-monitoring). For the purposes of this study, we treat them as a unit. The norm is for them to split the residual revenue evenly.

determine if their driver is driving recklessly on the road, while 32% report that they can rarely tell. Similarly, 80% of owners claim that they can only sometimes determine if their driver is responsible for damage to the vehicle, with 20% saying that they can rarely tell. Lastly, 56% of owners report that they can only sometimes tell if their driver is operating off-route, while 36% say that they can rarely tell.

This lack of visibility makes it difficult to effectively manage drivers and attribute poor business outcomes to their performance with certainty. Approximately, 75% of owners reported that they sometimes or always faced challenges managing their drivers, while 25% said that they rarely or never faced them. Similarly, 40% of the owners found it difficult or very difficult to manage their drivers, while another 40% had a neutral opinion about it (the remaining 20% of the owners reported that managing their drivers was either easy or very easy). Matatu owners describe many different challenges they could potentially address if they had complete visibility into their drivers' work, including drivers' dishonesty about revenue, target, fuel, vehicle location, required repairs, and police interactions, not working hard enough, showing up late, or not showing up at all, and not knowing how to interact with police. Most matatu owners have a negative view of the quality of drivers in the industry. Specifically, around 70% of the owners agreed that 50% or fewer of the matatu drivers in the industry are good or reliable. While owners believe that finding any driver should be quick (1 day on average), they recognized that finding a good and reliable driver is a much more challenging task that can take significantly longer (30 days on average). Moreover, 70% of owners believe their own drivers can improve their performance in some way. Hiring a new driver that will outperform an existing one is therefore difficult and risky because the quality of drivers is fairly low on average. In practice, this means that owners may choose to retain drivers they think are performing well-enough, even if they may not be entirely satisfied or convinced of their on-the-job performance.

The problems of moral hazard that firms face in this setting are not unique to this context. Although we are unable to determine how the exact magnitude of these problems compares across various contexts and industries, we have compiled evidence from various sources to illustrate that other firms encounter similar challenges. These include firms in other industries in Kenya (based on our own surveys), larger companies in the transportation sector, and other companies in different industries.

We sampled an additional 100 business owners (selling various goods such as electronics and clothes) in downtown Nairobi and Kisumu, and we find that they too are not able to fully and accurately observe their employees' behavior. Approximately 51% say they can only sometimes tell if an employee is not handling revenue correctly, 21% say they rarely can, and 21% say they never can. In terms of identifying unprofessional customer engagement, 63% of respondents say they can only sometimes tell, 13% say rarely, and 3% say they never can. Finally, 68% of business owners claim they can only sometimes tell if an employee is not handling the inventory correctly, 12% say they rarely can, and 2% say they never can.

The limited visibility business owners have contributes to the challenges they face managing their workforce. Approximately 60% report facing challenges managing their employees sometimes or always, while 40% indicate that they rarely or never encounter such difficulties. Similarly, 30% claim that it is difficult or very difficult to manage their employees, while 50% do not express a strong opinion, and 20% find it is easy. Owners identify many challenges they could potentially overcome if they had greater visibility into their employees' work. These challenges include employees not showing up to work or being late, not putting in enough effort to meet revenue targets, mishandling inventory, allowing unauthorized persons into the business premises, being dishonest about revenue and inventory (i.e., stealing), and failing to escalate significant issues. Many owners are relatively pessimistic about the quality of employees in the workforce. Approximately 40% of business owners believe that 50% or fewer workers in the industry are good or reliable. Moreover, while business owners claim that it should only take three days on average to find a new employee, they believe it will take ten times longer (30 days) to find a good/reliable employee. Some owners resort to hiring their family or close friends/connections because they trust them more. Finally, a majority of the owners (55%) think that their employees could perform better in their work and identify specific areas where they can improve, including how they interact with customers and manage inventory.

These surveys suggest that business owners in Kenya (matatu owners and general businesses) cannot observe many dimensions of their employee's behavior and struggle with employee management as a result. This finding aligns with data from other contexts, both within and beyond the transportation industry. de Rochambeau (2020) studies the trucking industry in Liberia where she documents how managers faced challenges in preventing drivers from engaging in behaviors that were harmful to the business, such as shirking and transporting unauthorized goods or passengers. According to a study by Startz et al. (2023) that focused on small businesses in Uganda, about 50% of the businesses surveyed encountered difficulties with their employees. These challenges included employees being late or absent, not putting in enough effort, having a poor attitude, not performing tasks well, and lacking trustworthiness. In Pakistan, Atkin et al. (2017) document that employees resist the adoption of technologies that are profitable for the firm by misinforming firm managers about the value of the technology. Finally, Caria and Falco (2022) find that entrepreneurs have low expectations of their employees, and they underestimate their workers' trustworthiness by 20%.

2.2 Monitoring

In 2015 companies started offering GPS tracking services in Kenya. Many insurance providers also began mandating that minibuses install GPS trackers for security reasons.[13] Monitoring capacity was still limited in the transportation industry however, which provided the ideal setting for documenting the costs and benefits of these technologies. Most medium and small-scale minibus owners had not installed them in their own vehicles at the time of the study because they were either prohibitively expensive (around \$600 per unit) or too complicated to operate. To fill this need, we worked with a Kenyan technology company to create a new monitoring system that was considerably cheaper and more flexible than other tracking systems. We describe the system in detail in Section 3.

The descriptive surveys we conducted in 2022 demonstrate that matatu owners perceive the importance of monitoring technologies and find value in using them. Approximately 50% of owners say that GPS tracking devices are always useful, while 40% of owners find them sometimes useful. Owners acknowledge that monitoring devices assist them in keeping track of their drivers' location, number of trips they made, distance travelled, and their speed. Furthermore, 50% of owners believe that having a GPS tracker would increase the likelihood of expanding their business. Finally, we find that half of the owners in our descriptive sample use GPS tracking devices in their vehicles for the same reasons we had designed the device that we offered to our experimental sample five years prior. These reasons include locating the driver, calculating the distance travelled, and monitoring instances of speeding, sharp-breaking, and sharp-turning.

There is also anecdotal and empirical evidence, primarily from high-income countries, that monitoring technologies can be useful in settings beyond the one we study. Although the effect sizes presented in our paper are specific to our context, the following discussion suggests that the advantages of monitoring may extend beyond our setting.

During our descriptive survey with general business owners in Nairobi and Kisumu, we discussed the importance of monitoring devices. We found that 30% of business owners

[13] 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.

consider them to always be important, while 33% say they are sometimes important. Interestingly, 40% of business owners said they would be more likely to expand their business if they had access to monitoring technologies. Some business owners have already invested in monitoring technologies, with CCTV cameras being the most accessible for small businesses. These cameras can provide information such as staff presence, business open and close times, customer traffic, and inventory management. However, more advanced software for digital inventory tracking, electronic sales records, and point-of-sale terminal usage is less easily accessible for small businesses.

There is also considerable anecdotal evidence from high-income countries that firms see value in monitoring. Dickens et al. (1989) document as early as 1989 that firms expend considerable resources trying to detect employee malfeasance. According to a study by the American Management Association, 80 percent of major companies in 2022 monitor the internet usage, phone and email of their employees. These data suggest that moral hazard remains a first-order concern that many firms in higher income countries have tried to resolve by investing in monitoring technologies. Interestingly, by combining data from the World Management Survey with the World Bank's GDP data catalog, we find that firms are more likely to utilize meaningful metrics for monitoring employee performance as the GDP of a country improves (Appendix Figure A.1). Although this analysis is purely correlational, it provides an indication that monitoring technologies are more prevalent in higher-income countries, where firms tend to be larger, and suggests that firms recognize moral hazard as a substantial problem worth investing in. Finally, an increasing number of studies suggest that both private companies and governments are adopting monitoring technologies, and the benefits of doing so are becoming evident (Hubbard, 2000, 2003; Gosnell et al., 2020; de Rochambeau, 2020; Liu et al., 2021; Gertler et al., 2023).

3 Experimental Design

3.1 Tracking Device and Software

To understand the impact of monitoring on the matatu industry, we developed the *Smart-Matatu* 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 physical tracking units were procured from a company in the United States (CalAmp). The tracking device has a GPS and gyroscope, which capture the vehicle's location and its vertical/lateral/forward and backward 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. Finally, an API call is generated each time the owner uses the app to request data from the server.

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 costly driving because owners can see if the driver is operating on roads that are known to damage vehicles. The second tab displays all the safety alerts that are captured by the device. The final tab conveys a summary of the driver's effort and safety. The effort section lists the total mileage covered and the duration the ignition was turned on that day. This provides treated owners with a more accurate measure of driver effort. Finally, the SmartMatatu app was also designed to collect daily information from both treatment and control owners, including the target; the amount the driver transferred; any repair costs incurred; and an overall satisfaction score for their driver's performance.

The technology cost approximately \$125, which reflects the cost of the device (\$85), the cost of installing the device (\$25), and the cost of storing/processing and delivering the data from the device through the app (\$15). This does not include the cost of maintaining the app, nor the cost of repairing the device. The maintenance of our app only required two software engineers (who were also charged with other project tasks), estimated to be around \$50,000 per year (\$25,000 for the six months of the study) according to our implementing partner. Consequently, we anticipate that the app maintenance costs could be kept relatively low and would decrease significantly as the number of clients increased. With only 300 matatus, the variable cost would amount to an additional \$83 per year or approximately \$6 per month. This is comparable to existing GPS tracker companies, who typically charge between \$5 and \$20 per month. We do not include the costs of repairs or replacing the hardware as there were fewer than five devices that needed to be replaced throughout the course of our study, and GPS tracker companies typically do not pay for the cost of repairing and replacing the hardware (beyond offering a one-year guarantee). Regardless of whether we consider the upper or lower bound of variable costs, the cost of providing our device is comparable to existing business training programs. McKenzie et al. (2021) reports that the cost per business of the training programs in the first wave of randomized experiments of business training programs varied between \$21 and \$740. More recently, Van Lieshout and Mehtha (2017) report the average cost for offering a 5-7 day course in 18 different countries to be approximately \$177.[14]

3.2 Treatment Assignment

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 drive 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. It took approximately four months to recruit enough participants across the nine major commuter routes (see Appendix Figure A.2). We successfully registered 255 owners, which we randomized into treatment (126) and control (129).[15]

We conducted installations and trainings from November 2016 to April 2017 (Figure A.9). The field team scheduled a time to meet each owner individually at a location of their choosing. The owner was compensated for the time their vehicle spent off-road to perform the installation of the device with a one-time payment of KES 5000 (\$50). We installed the

[14] We adopt the prevailing approach to pricing business development services, which involves considering pricing at or above *marginal cost* (Karlan and Valdivia, 2011; Drexler et al., 2014). This explicitly excludes the costs associated with developing the program. We believe the fixed cost of our initiative was also reasonably competitive. We hired one software engineer who charged \$100,000 to develop the prototype over three years (equivalent to \$30,000 per year).

[15] The recruitment process took time mainly because of the difficulties we faced scheduling meetings with owners, who have busy schedules and many other commitments throughout the day. While some owners remained hesitant to participate, their reasons align closely with the most common barriers to technology adoption identified in the literature. In particular owners did not want to take their matatu off the road for a day and forgo a day's revenue to install the device (they may be present-bias), while some were not confident they could navigate the app before the training even though the app was user-friendly (they may have incorrect priors). These reasons suggest a reluctance to invest time and money into learning new things, rather than a lack of enthusiasm for the product. More detailed information regarding enthusiasm for GPS technologies can be found in the context section.

trackers under the vehicle's dashboard to prevent tampering in both treatment and control vehicles. Our field team took both treatment and control owners aside and provided them with an Android smartphone with our SmartMatatu app pre-installed and trained them how to submit reports through the app. The app only allowed access to information for owners randomized into the treatment group, who received an additional 30 minutes of training on the features of the app's information section. We administered a short survey to the treatment 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 offered continued support to treatment owners to help navigate the app. Finally, we granted control owners access to the information from the tracker for two months at the end of the 6-month study period.

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. It was up to treatment owners to decide whether to reveal this information to their drivers.[16] 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. In other words, because control drivers knew about the device, the treatment effect identifies the impact of owners utilizing monitoring information rather than the impact of simply being observed. We believe this is the relevant margin to study, as the sustainability of monitoring technologies relies on owners effectively utilizing the information in some way.[17]

4 Data and Descriptive Statistics

4.1 Data Collection

We collect data from three different sources. Enumerators conducted in-person baseline and endline surveys. Next, we gather a panel of daily responses from owners and drivers through our SmartMatatu app and SMS surveys, respectively. Finally, the GPS tracker collects a

[16]Treated drivers learned what *types* of information the GPS could collect from owners directly as well. Note that since drivers were aware of the device's placement under the vehicle's dashboard, it is likely that they correctly assumed the trackers would not be able to monitor the precise amount of revenue collected from passenger fares nor the exact repair costs they paid to the mechanic.

[17]Since SACCOs generally include more than 500 members, our intervention would have affected less than 3% of the entire route, minimizing the risk of spillovers between treatment and control groups

wealth of data that we use to measure driving behavior.

We administer the baseline survey during the tracker installations. The owner baseline survey collects basic demographics, employment history, features of the matatu, and their relationship with the current driver. Similarly the driver baseline asks about driver demographics, experience as a driver, unemployment spells, and their relationship with the current owner.[18] We also use games to gauge drivers' risk aversion and driver/owner's propensity to trust one another (Sprenger, 2015). To measure risk, we ask respondents whether they prefer to receive KES 500 ($5) for certain or play a lottery to win KES 1500. The trust game is similar to Berg et al. (1995): we present owners with KES 500 and asked them to select an amount to be placed back in an envelope. This amount is then be tripled and delivered to a matatu driver who decides how much to keep for himself and how much to return to the owner. The amount the owner chooses to place in the envelope is recorded in the survey. At the end of the six-month period, we also conduct an endline survey focusing on business investment decisions. Finally, we run a willingness-to-pay experiment, offering owners two additional months of monitoring information through the app.

Next, we collect daily data from owners and drivers. For owners, we rely on our Smart-Matatu app, which provides a novel means of collecting high-frequency data in a challenging environment. Owners in the study are reminded daily via a notification on their phone to report on that day's business activities through a form located on the app. They are 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 is successfully submitted, owners received KES 40 via M-Pesa (a mobile phone-based money transfer service).

We collect similar information from drivers through SMS surveys (because the drivers are not provided with smartphones). Specifically, the message asks about whether the vehicle was on the road, the amount of revenue they collected, and the residual revenue they kept as a salary. We emphasize that all of the data they share remains confidential (in particular, it would not be shared with the owner) and they are compensated KES 20 for each submission. We check for differential reporting in revenue and salary as drivers in the treatment group might be concerned that we are sharing this information with the owner, who they know is gathering data from the app (Table A.1).[19] If drivers report lower revenue and salary, they

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

[19]Since the owner already knows whether the driver is working or not, there is no incentive for strategic misreporting across the treatment groups for whether the vehicle was on the road.

could try to give the impression that their earnings per mile are lower than they actually are, with the intention of persuading the owner to lower the revenue target. We regress revenue and salary on the number of miles traveled, an indicator for treatment, and an interaction term between the two. The coefficient on the interaction term is neither economically nor statistically significant, indicating that there is no significant difference in the relationship between mileage and reported revenue between the treatment and control groups. This reduces our concerns about differential reporting of revenue and salary.

Finally, we rely on the CalAmp tracking device data. 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.3 depicts the number of times vehicles licensed to one of the routes pass through a particular location. The figure illustrates that off-route driving is relatively common practice.

4.2 Descriptive Statistics

The owners in our sample are predominantly self-employed men in their late thirties (see Table 1). On average, they have been managing their own vehicles for four years and possess eight years of experience in the industry. The drivers in our sample are exclusively male, slightly younger than the owners, and have lower levels of education. On average, they have eight years of experience working as matatu drivers. The matatu vehicles themselves are primarily Japanese minibuses that have been imported and used for approximately thirteen years. Some of these vehicles come equipped with special features such as free wifi, sound systems, or TVs. The average purchase price for these matatus was approximately \$6,675. Appendix Figure A.4 shows a photograph of a matatu.

The owner sets a daily revenue target at baseline of approximately \$31 (KES 3,130), and they report receiving \$26 (KES 2,600) on average from the driver. The target amount set by owners exhibits some variability (standard deviation of KES 446), indicating that owners have some discretion in determining the desired target within industry norms. If owners have perfect visibility into their drivers' behavior or can learn about their drivers' performance over time, they may be able to trust that the driver will operate carefully on the road, and compensate them with a lower target. We explore this in Appendix Table A.2. The first column examines the relationship between the baseline target and owner-driver tenure, but we do not find a strong association, suggesting limited learning. In column 2, we include other driver characteristics, yet we do not observe any significant correlations emerging.[20]

In terms of driver performance, drivers report collecting approximately \$71 (KES 7,126) in passenger fares (revenue) throughout the day, and they retain approximately \$9.07 (KES 907) as their salary (Table 3). They spend an average of 14.8 hours on the road, covering a distance of 96.6 kilometers. The average daily repair costs hover around \$4.83 (KES 483). We might expect firm outcomes to differ for owners who know their drivers better or longer. Indeed, owners may be able to leverage their repeated relationship with drivers to learn how they operate and customize contracts accordingly, or by implementing effective monitoring mechanisms. To explore this possibility, Appendix Table A.3 examines the relationship between the four key business outcomes (hours ignition was on, repairs, revenue and profits) and 1) owner-driver tenure, 2) driver characteristics from the baseline survey and 3) driver risk aversion measured at baseline. We find no evidence that owner-driver pairs who have worked together longer have better outcomes. We find suggestive evidence that driver characteristics are correlated with core business outcomes, though not consistently: drivers who value their jobs more and have more experience invest slightly more effort, while drivers with more education have lower repair costs and higher profits. There is also suggestive evidence that drivers with lower risk aversion tend to have significantly higher repair costs (Appendix Table A.3, row 5). Since risk aversion is not observable to the owner, monitoring mechanisms could prove valuable in increasing firms' visibility into drivers' working styles, leading to potential improvements in overall business outcomes and performance.

Reckless driving is widespread within our sample (Appendix Figure A.5). We capture reckless driving in three ways: the share of days drivers exceed 75km/h, the share of days drivers are flagged for sharp-braking, and the number of hours drivers deviate from the designated route on average. The first two measures are based on research in the transportation safety literature that identifies speeding and sharp braking as significant predictors of unsafe driving. The last measure specifically focuses on off-route driving. We see that approximately 50% of drivers exceed 75km/h on more than 20% of the days they drive, and

[20] Variation in the baseline target seems to be most highly associated with differences in the quality of the matatu itself. The quality of a matatu affects the expected revenue the matatu will earn throughout the day. Column 3 demonstrates this by examining the relationship between the matatu's age, and number of features it has, and the target. We see that each additional year is associated with 17 fewer schillings (\$0.17), and each additional feature is associated with 100 more schillings (\$1). Finally, we explore the relationship between the target and the matatu's estimated market price (column 4), where we consider the market price to be a comprehensive indicator of the matatu's perceived quality. Here we see that a price increase of a thousand schillings (\$10) is associated with an increase in the target of 0.41 schillings (\$0.0041).

approximately 25% of drivers consistently exceed this speed limit over half the days they operate the minibus. It is worth noting that the speed limit in Nairobi is 50km/h. Next, we find that approximately 50% of drivers break sharply on 20% of the days they operate the vehicle, with around 25-30% of drivers braking sharply over half the days they operate the minibus. Finally, we observe that 50% of drivers spend approximately three hours per day traveling 400m beyond the designated route. While we can account for a maximum of 1.5 hours if the bus needs to be stationed beyond the route each day, this still means that 50% of drivers spend an additional 1.5 hours beyond their designated route throughout the day. These statistics illustrate the widespread nature of reckless driving and highlight the challenges that owners face in rectifying this behavior with their existing monitoring capabilities. Considering how prevalent these behaviors are, it could be difficult for owners to find drivers who perform better than their existing employee.

While some owners and drivers maintain long-lasting relationships, there is a substantial amount of turnover in the industry. The median duration of the working relationship between an owner and a driver is six months, and a quarter of the sample have worked with their current drivers for a period of three months or less.[21] The likelihood of an owner and driver separating (either through firing or quitting) in our sample is estimated to be approximately 0.1% per day in our sample. This implies that there is a 99.9% chance of the driver being rehired the following day, and the annual probability of driver-owner separation is $1 - 0.999^{365} = 30\%$. Being fired does not appear to impose substantial reputational costs on drivers. In the descriptive survey we conduct a few years later, over 80% of owners acknowledge that dismissed drivers can secure employment with another firm within the SACCO. Owners attribute this fact to a combination of factors: varying preferences among owners (70%), the necessity to settle for available options due to high demand and the scarcity of good drivers (70%), and the lack of information provided by drivers regarding if and why they were terminated (60%).

5 Model

On the basis of these descriptive facts, we now describe a contract model of the owner-driver relationship in the informal transit industry. The goal of this model is threefold. First, it allows us to precisely describe the mechanics that lead this type of contract to be inefficient.

[21]The average employment tenure of 14 months in Table 1 is heavily influenced by a small number of long-lasting relationships.

Second, we can derive predictions about the effect of monitoring on the driver, the owner, and firm outcomes. Finally, the model provides the basis for the structural estimation of driver and owner 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. 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 show how the resulting contract compares to a social-planner benchmark (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.

5.1 Setup

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 along two dimensions: effort that increases revenue with no costs to the vehicle (e.g. more hours on the road), denoted by e, and effort that increases revenue but damages the vehicle (e.g. reckless driving), which we simply call "risk" and denote by r.[22] He chooses (e, r), incurring disutility $\psi(e, r)$. On the basis of these actions, nature draws gross revenue $\tilde{y}$ from the revenue distribution $G(\cdot|e, r)$, which is assumed to be bounded from below by a subsistence income w. Revenue net of subsistence is $y = \tilde{y} - w \in [0, \bar{y}]$. Nature also draws repair costs $c \in [0, \bar{c}]$ from $F(\cdot|r)$. Repair costs depend on risk but not effort and accrue entirely to the owner. Conditional on

[22]Note that the use of "risk" is nonstandard in the literature, compared to, for example Ghatak and Pandey (2000). In our model, risk is 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 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.

effort and risk, the revenue and cost distributions are independent.[23]

The owner chooses whether to rehire the driver for the next day with some probability $p(\cdot)$ – the rehiring schedule. In the baseline environment without monitoring, this rehiring schedule depends only on the transfer: $p(t)$.[24] 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.[25] If the driver is fired he receives his outside option $\bar{u}$, and the owner pays a hiring cost h before drawing an identical driver.[26]

The timing of the game is as follows. At the beginning of the day, the owner and driver agree on the contract. The driver then makes driving choices (e, r) during the day. Based on (e, r), nature draws net-of-subsistence revenue y as well as repair cost c. The driver then transfers $t(y)$ to the owner and keeps a residual "salary" $y - t(y)$. Finally, the owner rehires the driver for the next day with probability $p(t)$, or $p(t, e, r)$ in case of monitoring. If he is rehired, the game repeats the following day; otherwise, he consumes his outside option $\bar{u}$ indefinitely, and the owner pays h to rehire an identical driver.

[23]See Appendix Section D.2 for functional forms assumptions of the technology and preferences.

[24]In Appendix D.4, we consider the possibility that owners may use information on repair costs as a signal of risk. Considering the noise in repair costs, we show that owners are limited in their ability to affect driver behavior by making rehiring contingent on repairs. This is consistent with our qualitative surveys which suggest that repair costs make for an unreliable signal of driver risk (80% of owners claim that they can only sometimes determine if their driver is responsible for damage to the vehicle, with 20% saying that they can rarely tell). This makes it difficult for owners to use repairs as an input in their firing decisions. Anecdotally owners reported that they only used repair costs to fire drivers under extreme cases involving large accidents. As a result, we choose to exclude repair costs from the firing function as this improves the tractability of the model and better matches the situation on the ground.

[25]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).

[26]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. In any case, less than 15% of owners get a new driver over the course of the study, limiting the impact of adverse selection in our study period.

5.2 Baseline Contract Without Monitoring

In the status quo contracting problem, the owner maximizes the sum of expected transfers and the continuation value, minus the cost of risk and the expected cost of firing:

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

subject to

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

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 constraint is the participation constraint, which restricts driver utility to be at least as great as his outside option. 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 second constraint is the incentive compatibility constraint, which requires that the driver choose the level of effort and risk that maximizes his utility. 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, balancing his take-home pay against the probability of rehiring.[27]

[27] 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.

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 uncertain revenue.

Optimal rehiring and transfer schedules. We define the rehiring schedule using a the daily target T, which defines the level of transfers above which re-employment is guaranteed. Under the assumption that the owner prefers less risk than the driver, we can solve for the optimal transfer and rehiring schedules:[28]

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

$$p(t) = 1 - \frac{T - t}{\delta U - \bar{u}}$$

for all t such that $p(t) \geq 0$ and zero otherwise; and $p(t) = 1$ for all $t \geq T$. That is, 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, which we illustrate in Figure 2. 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 the change in the rehiring probability times the discounted value of that relationship $p'(t)(\delta U - \bar{u})$) exceeds the direct value of keeping that dollar (which is just 1). This implies the slope of the rehiring schedule needs to surpass the inverse of the discounted value of the relationship, $1/(\delta U - \bar{u})$.

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 driver's 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 $1/(\delta U - \bar{u})$, the driver's utility simplifies to $\mathbb{E}[y|e, r] + \delta U - T - \psi(e, r)$. The driver optimizes over this expression and chooses (e, r) to equalize the marginal revenue of effort and risk to the marginal cost. We call this the driver bliss point because it is the choice of effort and risk the driver would make if they

[28]See Appendix Section D.2 for more discussion about the justification and implications of this assumption.

were operating the bus on their own without any consideration for repair costs. The owner *could* induce higher effort-risk bundles by setting a rehiring schedule that is steeper than $1/(\delta U - \bar{u})$. 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 $1/(\delta U - \bar{u})$ because the driver would keep the marginal dollar rather than transfer it, without actually lowering effort and risk. Because we assume the owner prefers less risk than the driver (Panel B of Figure 2), she contents herself with the lower bound of risk induced by the minimal slope $1/(\delta U - \bar{u})$ 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.[29]

Inefficiency of baseline contract. We assess the efficiency of the baseline contract by comparing it to the optimal decision of an integrated decision maker (or social planner), taking into account both the repair costs due to risk as well as the disutility of effort and risk. In Appendix D.3.2, we show that the baseline contract is inefficient compared to the social planner's solution due to excessive risk taking by the driver. Risk is oversupplied relative to the social optimum because the driver is not accounting for repair costs accruing to 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.[30]

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.

5.3 Introducing Monitoring

With monitoring, the owner now observes the driver's effort and risk choices and conditions her rehiring schedule on them, in addition to the transfer: $p(t, e, r)$ instead of $p(t)$. Therefore, the solution to rehiring schedule solution becomes

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

[30] If effort and risk are weakly substitutable then higher risk could induce higher effort than the social optimum level of effort e^*.

$$p(t,e,r) = \begin{cases} 1 - \frac{T-t}{\delta U_M - \bar{u}} & \text{if } e = e_M \text{ and } r = r_M \\ 0 & \text{otherwise} \end{cases}$$

where (e_M, r_M) is the (owner mandated) effort-risk choice under monitoring. 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. Because the owner still has to rely on a target contract, she chooses an effort-risk profile under monitoring (e_M, r_M) that balances the size of the expected transfer and the expected repair costs.

Predicted effects of monitoring. This result yields several predictions for how key outcomes will change under monitoring (a proof is in Appendix D.3.2):

1. **Effort will increase and risk will decrease:** Compared to the baseline contract, the owner can now explicitly contract on higher effort provision ($e_M > e_B$) and lower risk ($r_M < r_B$) moving the driver to a more profitable mix.[31]

2. **Revenue may rise or fall:** The effect on revenue is ambiguous. The owner could settle on an effort-risk bundle that yields lower expected revenue if it also yields a larger drop in expected repair costs.

3. **Profits increase:** Profits will unambiguously increase due to lower repair costs.

4. **The targets falls if revenue falls:** If the revenue collected by the driver falls, the optimal target will also fall as the owner needs to compensate the driver for lost salary. Note that falling revenue is sufficient but not necessary for the target to fall.[32]

[31] Here, e_B is the amount of effort the driver supplies at his bliss point (which is what the driver would choose without monitoring).

[32] There are two forces that influence the owner's decision to re-optimize the target. First, the driver is worse off from having to adopt a new effort-risk bundle that differs from the one he previously selected. This increases the risk that the driver does not make any transfer at all as he is less concerned about losing his job. The owner needs to compensate the driver for this loss by lowering the target, thereby increasing the value of the job. Second, the owner will respond to a change in revenue. If expected revenue rises, the owner will increase the target in an effort to capture some of this surplus. However, if expected revenue falls, this reinforces downward pressure on the target. If revenue collected stays the same, we expect the target will still fall as the owner needs to compensate the driver for larger disutility from work. 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.

5. **Ambiguous welfare effect**: Finally, we show that 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.[33] We return to these calculations in Section 7.

6 Experimental Results

We now discuss the reduced-form impacts of the experiment. We fist provide evidence for basic features of our contract model, and we then turn to a discussion of the impact of the intervention.

6.1 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: 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 subsistence income. Second, the figure also shows that owners' satisfaction with their driver increases with the size of the transfer, as suggested by the rehiring schedule.

6.2 Treatment Take-Up

We investigate the degree to which owners engaged with the monitoring app and how the app affected self-reported management practices. We monitor owners' usage of the device

[33]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 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 $\psi(e_m, r_m)$, 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.

by tracking the API calls that are generated every time the owner logs into the app and requests different pieces of information. We find that 70% of owners consult the app weekly, while 50% use it daily (Appendix Figure A.6 Panel A and D). Since the app also included the survey we asked owners to complete, we can evaluate their usage of the app separately from the survey in two different ways. First, we can look at API calls to the main information dashboard of the app, which is distinct from the survey form. Panel B reports weekly use of the main dashboard, and we can see that the usage rate closely tracks overall app use - leveling off at around 70% each week by the end of the experiment. Second, we can examine the average number of hours the owners were active on the app. Panel C shows that owners use the app for eight to nine hours per week by month six. This is well beyond the time it takes to complete the 3-minute survey, suggesting significant engagement with the app.

We also confirm that owners are internalizing the information we provided through the app. At endline, we asked owners to state whether they knew the revenue earned, the number of kilometers driven, and the extent of off-road driving on the most recent day their vehicle was active (Table 2). We find that owners in the treatment group are 27 percentage points more likely to know about the number of kilometers driven and 45 percentage points more likely to know about the instances of off-route driving (columns 1 and 2). 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 (column 3).

We also see that treated owners find it easier to monitor their drivers than control owners, and spend less time monitoring their drivers. We ask owners at endline to rate how challenging it is to monitor their employees on a scale from 1 (not hard) to 5 (very hard). Having access to the information reduces the reported difficulty level by just under 2 points (Table 2, column 4). In other words, control owners maintain that monitoring is hard while treatment owners reveal that it is easy. Furthermore, we ask owners whether the amount of time they spend monitoring has increased or not in the last six months. We see that 72% of treated owners report a decrease in the time they spend monitoring (Table 2, column 5).

6.3 Results

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

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

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;[34] and ε_{id} is an error term, which we allow to be arbitrarily correlated within i across days. We cluster the standard errors at the matatu/owner/driver level. This design allows us to examine the treatment effect as it evolves over the six months of the study.[35] This is important because it took a few months for owners to become comfortable with all the features of the monitoring app. We consider all the impacts presented below as intent-to-treat (ITT) estimates, as some owners did not consistently use the app.[36]

Effort. We proxy driver effort by the number of hours the matatu is operating and find that operating hours increase by one hour per day on average by the third month after installation (p-value = 0.072) and rise steadily until the end of the study (Figure 4 or Table 3, column 1 and 2).[37] By month six, effort levels increase by 1.45 hours per day on average in the treatment group, a 9.8% increase in drivers' labor supply (p-value = 0.055). 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. However, there is no significant increase in safety-related outcomes. With more hours on the road, we also see the number of kilometers increase by 13 kilometers per day on average (13%) by month six (p-value = 0.062), which corresponds to an extra trip to or from the city center.

[34] Specifically, we include as control variables 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.

[35] Our trial was registered on the AEA RCT Registry (# AEARCTR-0001482). After registration, we split the analysis of the randomized control trial into two papers, one focusing on firms, and the other focusing on road safety. The major deviations from this document include only reporting firm level outcomes for this paper (please see our companion paper (Lane et al., 2022) for safety outcomes), dropping a control group that does not receive a device because it was not feasible to implement in the field, and presenting the main specification by month because of the important learning dynamics. We do however present the specification detailed in the registry which pools all months of our data together in the first panel of Appendix Table A.4. More broadly, our approach to this study has been to focus our primary analysis on a parsimonious set of outcomes, each derived from theory (effort, costs, target, revenue, profit, salary, growth). Following the guidance of Banerjee et al. (2020), our readers may wish to interpret any reported analysis outside of these primary outcomes as a secondary analysis.

[36] Table 3 presents all the treatment effects by month in table form, and corresponds one-for-one to the figures presented below. The second panel of Table A.4 presents treatment effects from a pooled regression of the same outcomes on an indicator for being treated in the first three months, being treated in the last three months of the study (panel B), and the standard controls and fixed effects as in Equation 2. This provides another way to showcase the contrast between early and later stages of learning.

[37] The monitoring device powers on and off with the matatu, so we can track when the vehicle is operating.

Risk. Treatment drivers also appear to take substantially less risk. 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 (p-value = 0.022). Next, we investigate whether this change in the distance from the licensed 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.[38] Taking fewer bumpy roads that may damage the vehicle is visible in the acceleration data. Lateral acceleration measures tilting from side to side, while vertical acceleration captures movement upwards and downwards. We find that the distributions of lateral and vertical acceleration in the treatment group tighten around zero, consistent with a reduction in reckless and damaging driving. We can reject equality of treatment and control distributions by applying a K-S test, which returns a p-value below 0.001 for both measures of acceleration.

These findings are consistent with anecdotal evidence that one of the greatest sources of risky driving is to 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. When owners have access to the GPS technology, they can monitor where drivers are at any point throughout the day and mandate they stay on the designated routes.

Repair costs. In line with model predictions, these changes in driver behavior translate into lower repair costs. Figure 6 (Table 3, column 3) shows that repair costs reported by treatment owners decline steadily relative to control owners. By the third month, daily repair costs for treatment owners fall by KES 124 ($1.24) (p-value = 0.12). By the sixth month, daily repair costs are KES 216 ($2.16) per day lower for treatment owners (p-value = 0.037). The magnitude of the effect is large: these reductions represent a 44.6% decrease in daily repair costs. They represent a major business expense for owners, which makes the impact of the monitoring technology significant.

It is also important to rule out any alternative explanations for these effects on repair costs. Specifically, it could be the case that drivers inflate repair costs, and the device reduces

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

their incentive to do so because they are more likely to be caught. This is unlikely to be the case for larger repairs, however, because they are incurred directly by the owner and/or will be validated with the mechanic. We create an indicator for whether repair costs exceed $10 (1000 KES – 80th percentile). The probability of incurring a large repair cost decreases significantly in month 6 (7.71 percentage points, p-value = 0.042) (Appendix Figure A.7, and Table 3 column 4). This implies that the decrease in the repair costs that we observe cannot be entirely driven by inflated repair costs.

These results are in line with the model predictions 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 choices.

Revenue. The top panel of Figure 7 (Table 3, column 5) shows that the effect on revenue is close to zero, and may be declining slightly. 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.

Target. While imprecisely estimated, there is some indication that treated owners set a slightly lower target than control owners. Figure 7 (Table 3 column 6) shows that by month six, the daily target amount is 149 KES ($1.49) below the control group, representing a 4.9% decrease (0.2 standard deviation). The effect is not statistically significant (p-value = 0.114), but a downward trend is visible and suggests that the information may allow managers to re-optimize the terms of their employees' contracts. We also see that drivers are more likely to make the target on occasion. Appendix Figure A.8 (Table 3 column 7) shows that the probability of making the target increases in some months. Drivers are 7% more likely to make the target in the third month (p-value = 0.142), 13% more likely in the fourth month (p-value = 0.014), and 5% more likely by month six (p-value = 0.317).

The fact that owners retain the target structure even under monitoring is consistent with our model predictions. They will now base their decision about whether or not to rehire the driver on the transfer the driver provides, as well as the effort and risk the driver supplies. A wage contract is not feasible because the tracking device does not reveal information about revenue, such that the owner must continue to provide incentives to the driver to make transfers. Similarly, a fixed rent contract is still infeasible in this context because limited liability prevents the driver from paying a rental price upfront. Our model also states that if revenue falls under the newly contracted effort-risk bundle, the owner should set a lower target to compensate the driver.

Profits. We now turn to investigating the impact of the monitoring device on firm performance. Specifically, we are interested in determining whether the information we supplied allows companies to generate higher profits and ultimately expand their operations by adding more vehicles to their fleet. Firm profits are measured by subtracting costs (repairs and driver salary) from total revenue. With revenues staying largely the same and repair costs falling significantly, we would expect profits to rise. Figure 8 (Table 3, column 8) shows that daily profits rise by approximately 13% in month four and five (449 KES ($4.49), and 453 KES ($4.53) per day, respectively) for treatment owners (p-value = 0.046 and 0.035, respectively).[39]

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 approximately $125, an amount that could be recovered in less than three months.[40] 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 8.

The devices' return on investment suggest they are likely to be a worthwhile investment for owners in the short and long run. One of the reasons we do not see more matatu owners adopting them is because they did not exist in this form on the market at the time of our study. The options were either much more expensive (approximately $600 and monthly installments), or had more limited capacity. Without having tested their efficacy, owners were hesitant to make the investment, consistent with classic work on technology adoption (Foster and Rosenzweig, 1995). It is also worth mentioning that our profit gains are in line with some of the more successful business training programs documented in the literature. The cost of these trainings range from $21 to $740 and last a few weeks at most (Bloom et al., 2013; McKenzie and Woodruff, 2017; Berge et al., 2015; de Mel et al., 2014; Valdivia, 2015). Our technology has the added benefit of requiring a single up-front payment for continued use. Moreover, it requires relatively little coordination and training.

[39]Profits fall in month 6 to 179.75 KES ($ 1.80) (p-value = 0.430).

[40]If we include variable costs under the conservative assumption that a company could only ever manage 300 devices at a time, the cost of the device increases to $208, an amount that could be recovered in less than four months.

We also explore quantile treatment effects for these core business outcomes (effort, repair costs, revenue, target, profit) at the 10th, 25th, 50th, 75th, and 90th percentile of these distributions (Tables A.5 to A.9). Driver effort, as measured by the number of hours the ignition was on, improves on most days. The impacts on reckless driving can also be detected across the distribution, albeit most pronounced for days with very large repair costs. This suggests that the monitoring device contributes to a decrease in the necessity for significant repairs, rather than reducing minor and frequent maintenance tasks. Furthermore, the most significant impacts on revenue are observed at the 90th percentile, perhaps suggesting that mitigating driver risk-taking prevents drivers from reaching those highest revenue days. Finally, the positive impacts on profits are concentrated at the 10th and 25th percentile of the distribution, likely stemming from a reduction in larger repairs. These quantile treatment effects suggest that treatment owners use the device to curtail their drivers' reckless behavior, which reduces the probability of very large repairs (and potentially limits the number of very good revenue days). In other words, the monitoring device improves outcomes for any firm who experiences these bad days by lowering the probability of reckless behavior. We see that the median owner experiences these large repairs 12% of days their vehicle is on the road, and 90% of owners incur these costs at least once over the course of the study. While reducing reckless driving may lead to a decrease in the number of higher revenue days, it ultimately improves firm profitability on average, which aligns with the business objectives of the owners. Finally, we see that the impact on the target is largest and statistically significant at the 10th percentile of the distribution, which means some owners are lowering their target below industry norms, rather than owners with high targets re-adjusting.[41]

Firm growth. We find weak statistical evidence that treatment firms are more likely to grow their business than control firms. We measure firm growth by the number of vehicles that owners have in their fleet at endline. We find that treatment owners have 0.129 more vehicles in their fleet on average than control owners (Table 4), a ten percent increase in

[41]We also explore two dimensions of heterogeneity. First, span of control models suggest that better monitoring can help more productive firms most. While we do not have measures of baseline productivity, we investigate heterogeneous treatment effects by owner education and raven's score, as they may be correlated with firm productivity. Second, if owners had visibility into drivers' operations and could discipline drivers' behavior over time by learning how they operate and customizing contracts accordingly they may be able to achieve better outcomes. We explore heterogeneous treatment effects along three margins that relate to how much visibility owners have into driver types (or lack thereof). This includes 1) owner-driver tenure, 2) an observable driver characteristic (driver education) and 3) one unobservable characteristic that directly relates to drivers' inclination towards reckless behavior (risk aversion). Overall, there are no clear patterns that emerge from the data on any of these margins.

fleet size (p-value = 0.091). While not statistically significant, Table 4 also demonstrates that owners invested in the interior of their vehicles through the purchase of items such as higher quality seating, lighting, and sound systems (p-value = 0.19).

There are a number of reasons why the monitoring device could have encouraged treatment owners to grow their businesses more actively (even if they did not fit a monitoring device in this second vehicle). First, these effects could be driven by the cost savings and profit gains that owners reap, which may make it easier to take a loan for a second bus. Second, because the owners operate both buses on the same route, the information gathered from the monitoring device in the first bus could provide insights into the operations of the second bus as well, making it easier to manage. For instance, knowing the number of trips completed by the first matatu can provide information about the overall traffic conditions the second matatu faced as well.[42]Third, owners also report that managing their vehicles has become easier, which could lower the mental burden of taking on a second bus (Table 2, column 4 and 5). Finally, owner's perceptions of their drivers' performance has improved (Table 5). Treatment owners' assessment of whether their drivers' skills improved increases by 0.63 points (where they could be assigned a -1 for worse driving, 0 for no change, and 1 for better driving) (p-value = 0.00). Owners also report that drivers are significantly more honest, and they trust their drivers more (columns 1 and 2). These broad performance improvements could make the prospect of expanding to a second bus less daunting.

Salary. Finally we consider the monetary gains to drivers. Figure 9 shows that the impacts on driver salary per hour are close to zero. While not a formal prediction of the model, the effect of the tracking device on driver take-home pay is ambiguous. The impact will depend on how revenue changes, and how much the owner adjusts the target. However, it is important to note that this is not the only metric that informs driver welfare. We need to consider how changes in driving behavior and the relationship with the owner affect the driver, which we discuss in the next section.

Separations. We do not see any differences in the rate of separation between drivers and owners in the sample (Appendix Table A.10). The rate of separation in the control group is 0.19 and the difference between treatment and control is small (0.03) and not statistically significant. Nevertheless, we test whether new drivers are responsible for the

[42]In the descriptive survey exercise we conducted in 2022, 50% of owners with multiple buses said that having only one GPS device in one of their minibuses would be sufficient to help them understand the behavior of other buses in their fleet.

observed effects by excluding days when new drivers operate the vehicles, and by comparing new drivers across treatment and control groups. We find the results are qualitatively similar to those we obtain from the full sample, and that new drivers are similar across treatment and control groups. This suggests that driver selection is not driving the effects we observe in the experiment.

7 Structural Estimation

The previous section provides reduced-form evidence for the effect of monitoring on driver behavior, firm outcomes, and contract parameters. We now proceed to estimate the model laid out in Section 5, with two goals in mind. First, we seek to quantify the value of the target contract at baseline to both the owner and the driver. These valuations provide a basis to assess the welfare impact of target contracts, which are common in informal transit. Second, as firms adopt monitoring technologies more widely, it is important to understand their distributional consequences. Hence, we are interested in how the introduction of monitoring changes these valuations and the corresponding welfare consequences.

7.1 Estimation Procedure

The intuition for the estimation procedure is that we seek to match contract characteristics observed in the data to the corresponding predictions by the model. These contract characteristics include the rehiring rate, the daily target, and the driver's valuation of the contract. The model predictions of these moments are based on characteristics of the production environment (i.e. revenue and repair costs) rather than the contract characteristics themselves. These predictions are also a function of a set of unobserved parameters that we estimate. In the estimation, the parameters of the model adjust so as to make these observed (i.e. reduced-form) and predicted (i.e. structural) contract characteristics as similar as possible. All details on the estimation procedure are in Appendix E.

Identification. At baseline (i.e. without monitoring), there are three parameters in the model: the driver's outside option, $\bar{u}$; his disutility from his chosen effort and risk at baseline, $\psi_B = \psi(e_B, r_B)$; and owner firing costs, h. These three parameters are identified by three moments, each of which consists of a structural component derived from the model, as well as a reduced-form component that we can observe in the data across owner-driver pairs indexed by i.

The first moment is the expected rehiring rate, whose structural component is given by

$$\mathbb{E}\left[p(t)|e_B, r_B\right] = 1 - \frac{G(T|e_B, r_B)\left(T - \mathbb{E}\left[y|e_B, r_B, y < T\right]\right)}{\delta U - \bar{u}}$$

under the optimal target contract without monitoring. Its empirical equivalent in the data, i.e. the average rehiring rate, is denoted as p_i.

The second moment is the driver's daily valuation of the contract at baseline, which is given by

$$U = \frac{\mathbb{E}\left[y|e_B, r_B\right] - T - \psi_B}{1 - \delta}$$

and its empirical equivalent in the data, i.e. the driver's stated value of the contract, is denoted by U_i.

Finally, the third moment is the owner's target choice so as to maximize the value of her business:

$$T = \arg\max_{\tilde{T} \geq 0} V(\tilde{T})$$

with

$$(1-\delta)V(T) = T - G(T|e, r)\left(T - \mathbb{E}\left[y|e_B, r_B, y < T\right]\right)\left(1 + \frac{h}{\delta U - \bar{u}}\right) - \mathbb{E}\left[c|r_B\right].$$

and its empirical equivalent, i.e. the observed target, is denoted by T_i. Each of these three expressions follow directly from the model after imposing the optimal transfer and rehiring schedules.

As we now describe, each parameter is identified off of all three moments. They are separately identified by the strength and direction with which they affect the moments. We can see that the driver's outside option, $\bar{u}$, is directly identified in the moments $\mathbb{E}\left[p(t)|e_B, r_B\right]$ as well as T, and indirectly in U through the appearance of T. The intuition is that the driver's outside option affects the minimal rehiring slope required to provide transfer incentives, which in turn affects his rehiring probability and his daily target. Specifically, as the driver's outside option improves, the marginal benefit of making a transfer must increase for the driver to want to make the transfer. This means the marginal change in the probability of rehiring with respect to transfers must increase. A steeper rehiring function means the optimal target falls, but the overall probability of being fired increases. These changes then indirectly affect the driver's contract valuation as well. Simply put, a better outside option forces the owner to offer a better deal: the driver has to reach a lower target, which increases the value of the contract to the driver, even though he accepts a higher risk of getting fired.

Next, the driver's disutility ψ_B is directly identified in U, but also indirectly in the other two moments as U appears in those moments as well. The intuition is that driver disutility affects the value of the contract, which in turn affects the minimal rehiring slope required for transfer incentives (and subsequently the rehiring probability and the target). That is, if the driving job is tougher then it becomes less appealing, and the owner needs to sweeten the deal by offering more lenient contract terms.

Finally, the hiring cost h is directly identified in T and indirectly in the other two moments through T. Firing costs affect the owner's contract valuation and her optimal target, which in turn affects the driver's contract value and the rehiring rate. Intuitively, if firing a driver puts a larger burden on the business, the owner seeks to lower the probability of firing by lowering the target, which increases the driver's contract valuation.

With the introduction of monitoring, we have to estimate one additional parameter, which is the driver disutility under the monitoring contract, which may change from ψ_B to $\psi_M = \psi(e_M, r_M)$. To identify this parameter, we use the two moments we can identify in the treatment group (expected rehiring rate and the target). The intuition is the same as before: higher driving disutility requires a better deal for the driver, which lowers the target and increases the rehiring rate.

Calibration of additional parameters. We need to calibrate two additional parameters for which we lack moments to identify them separately. Specifically, this applies to the discount factor δ and the subsistence income w. We ensure that estimation results are similar if we vary these calibrated parameters in the appendix.

Estimation. We estimate our parameters of interest via generalized method of moments (GMM), minimizing the distance between structural and reduced-form components. We assume that the revenue distribution $G(\cdot)$ is Normal since it simplifies the estimation of average revenues below the optimal target, $\mathbb{E}[Y_i|Y_i < T]$. Our data $\mathbf{X}_i$ consists of rehiring rates p_i, driver baseline valuations U_i, and target T_i—which make up the reduced form components—as well as revenue Y_i with standard deviation s_i, and repair costs c_i—which appear in the structural component. Our GMM estimator for our control group sample of size N_c minimizes $\left(N_c^{-1}\sum_{i=1}^{N_c} m(\mathbf{X}_i,\theta)\right)' \mathbf{W} \left(N_c^{-1}\sum_{i=1}^{N_c} m(\mathbf{X}_i,\theta)\right)$ where

$$m(\mathbf{X}_i,\theta) = \underbrace{\begin{pmatrix} p_i \\ U_i \\ T_i \end{pmatrix}}_{\text{Reduced form}} - \underbrace{\begin{pmatrix} \left[1 - \frac{1}{\delta U - \bar{u}}\left\{G(T)\left(T - \mathbb{E}\left[Y_i | Y_i < T\right]\right)\right\}\right] \\ \frac{1}{1-\delta}\left\{\mathbb{E}\left[Y_i\right] - T - \psi_B\right\} \\ \arg\max_{T\geq 0} \frac{1}{1-\delta}\left\{T - G(T)\left(T - \mathbb{E}\left[Y_i | Y_i < T\right]\right)\left(1 + \frac{h}{\delta U - \bar{u}}\right) - c_i\right\} \end{pmatrix}}_{\text{Structural}}$$

is the vector of moments; T is the optimal target according to the model and the production environment (i.e. revenue and costs); and $\mathbf{W}$ is a weighting matrix consisting of the inverse variance of the moments. Since we are interested in conditions at baseline, we use data from the control group only.

Estimation with monitoring. After the introduction of monitoring, we assume that the outside option $\bar{u}$ and the firing cost h are unchanged, but the driver's disutility under monitoring ψ_M may be different from baseline disutility ψ_B. To estimate this additional parameter, we can again use the expected rehiring rate $\mathbb{E}\left[p(t)|e_M, r_M\right]$ and the optimal target T, but we can no longer use the driver's contract valuation U, which was only measured at baseline. Hence, we now estimate a very similar GMM system, but this time with only two moments and one parameter; and using data from the treatment group instead of the control group.

Targeted and untargeted moments. In addition to the targeted moments described above, we also observe untargeted moments which we do not match in the estimation procedure, including driver salary and owner profits.[43] In both estimations (baseline or monitoring), we can evaluate how well the model-predicted structural components can accommodate the empirical components for both the targeted moments and untargeted moments. The extent to which the targeted moments exhibit no gap between model prediction and data reflects the model's capacity to match different aspects of the data through the lens of the model's optimal contract. The size of the gap in untargeted moments provides additional validation for the model being a good fit for the data.

[43] While these are separate moments in the data, their structural components are linear functions of targeted moments, so they do not offer linearly independent variation that could be used to pin down the calibrated parameters.

7.2 Status Quo Valuation

Table 6 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 $2.47 (SE $0.70), while we estimate firing costs to be equivalent to $263 (SE $19), roughly equivalent to 11 days lost profit. The driver's outside option is estimated to be $8.57 (SE $0.37), or $1.57 above subsistence, which is approximately equal to the average unskilled daily wage in Nairobi.[44]

The model succeeds in matching the observed moments in the data. Expected firing and the target are matched nearly exactly, while the driver contract value differs by $0.70, a small difference given the standard error of 71.6. Furthermore, the untargeted moments also match reasonably well. Predicted driver salary is $0.80 above observed salary while predicted owner profits are $0.80 below the observed level, although neither difference is statistically significant.[45] Overall, the model does well in matching both the targeted and untargeted moments, which we believe suggests that the model is a good fit for the data.

Finally, we can use the model to estimate owner contract value and the total welfare of the contract under the status quo. The model estimates that the contract confers substantial value to the owner, at $2,177 (SE $10). Adding this value to the estimated driver contract value of $507 (SE $70), we estimate that the total welfare accruing to both the owner and driver is $2,684 (SE $71). This implies that despite the unobservability of revenue and driver actions, the owner is still able to capture approximately 80% of the total value generated by the business via the use of the target contract. However, the contract still provides substantial value to the driver above their outside option, consistent with the view among drivers that their job is "good".

[44] Appendix Figures A.11 to A.13 investigate the sensitivity of the model predictions to different values of the parameters. These Figures plot how the three structural moments (rehiring probability, target, and driver value) and three un-matched moments (driver salary, owner profits, and owner value) change as the three parameter estimates $\bar{u}$, $\psi(e, r)$, and h vary.

[45] These moments are linearly dependent, so the gap between observed and predicted salary will always be the same magnitude as the gap between observed and predicted profits. One un-modeled explanation for this remaining gap is that it is not always true that the driver gives the entire revenue amount to the owner when it is below the target. This may happen because idiosyncratic events increase drivers' need for cash on a given day (e.g. a relative became sick and needs to go to a hospital). Therefore, drivers may deem the value of lying and keeping some revenue for themselves worth the risk of being fired when on other days they would not. In general, this will lead to a higher reported salary (and lower reported profit) in the data than the model would predict. However, we choose not to model this behavior because its stochastic and unobservable nature would complicate model with little gain in economic insight.

7.3 Valuation with Monitoring

Our second exercise shown in Table 7 aims to estimate how the introduction of monitoring affects owner and driver welfare. As discussed above, we hold the driver's outside option and the owner's firing costs fixed from the status quo estimation. The final unknown model parameter is driver disutility under monitoring, $\psi(e_M, r_M)$. In Panel B of Table 7 we show its estimated value from the GMM procedure is \$3.74 (SE \$1.74), which is a \$1.27 or 52% increase from the disutility estimated under the status quo.

In Panel C, we report the observed changes to moments calculated by comparing the treatment and control data, the structural model predicted changes in these moments, and finally the difference between these estimates. Starting with the targeted moments, the structurally estimated treatment effects closely matches the observed changes in the firing probability and target. There is a negligible difference in the firing probability and a \$0.15 difference (SE \$1.30) in the predicted treatment effect on the optimal target. For the untargeted but observed moments of driver salary and owner profits, the structural estimates do slightly less well, but still within reasonable bounds. The structurally estimated change in salary of \$0.9 is larger than the \$0.1 we observe in the data and the model prediction change in profits of \$0.8 is \$0.9 smaller than the observed \$1.7 reduced form change in owner profits, although neither difference is statically significant. In sum, while matching well overall, the model predictions slightly underestimate the benefits of monitoring to the owner and overestimate the benefits to the driver.[46]

Finally, this exercise estimates the changes in driver and owner welfare. From the theoretical model, we have a clear prediction 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. If driver disutility increases only slightly or falls, driver welfare *rises* along with the owner's. Despite being a pareto improvement, this outcome is 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.

[46] One explanation for this under-estimation of owner benefits is that the monitoring device may allow owners to reduce (un-modeled) driver behavior where they do not give the full revenue amount when it falls below the target (as discussed in the footnote above). Treated drivers may expect their owners to have a better signal about their revenue based on distance driven which could prompt them to transfer more to the owner, thereby lowering the driver's salary and increasing the owner's profits by more than the model predicts.

Figure 10 Panel (b) illustrates these dynamics for different costs. When driver disutility increases by less than roughly 20%, the driver would be better off with monitoring than without. Above this level, the owner's gain from monitoring comes at the expense of driver losses. Our actual point estimate reported in Table 7, Panel C for changes in the driver's contract value is -\$19.6 (SE \$26), consistent with the 52% estimated increase in ψ. Meanwhile, we estimate that owner contract value increased by \$83 (SE \$125), in sum leading to a total welfare increase of \$63 (SE \$128). These structural estimation results suggest that monitoring leads to small efficiency gains, with some redistribution from the driver to the owner. It is worth highlighting that these point estimates are imprecisely estimated and we cannot rule out larger gains (or losses) for the owner and driver.[47]

To evaluate the plausibility of the owner's valuation of the monitoring device, we can compare our estimates to our findings from a 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 \$45 as compared to the model estimate of \$83 (SE \$125). This suggests owners do perceive the monitoring devices to be valuable and their willingness to pay is broadly consistent with the structural model.[48]

8 Discussion of Welfare

While we find that owners end up offering a welfare-reducing contract to drivers in this case, this did not have to be the case as we discuss above. Moreover, there is an important caveat to this finding. These welfare estimates do not account for changes in the intangible relationship between owners and drivers. The relationship between owners and drivers is notoriously fraught with mistrust: under the status quo, owners often suspect that drivers are cheating them and driving the matatu recklessly; conversely, drivers often complain that owners second-guess their reports, and refuse to give them the benefit of the doubt when things go awry. In focus groups with drivers during the development of the monitoring

[47] Note that the reduced form estimate for the effect of monitoring on profits in section 6 (\$1200) is not discounted to present value. If we apply the same discount rate as we do in the structural model, we estimate the present discounted value of the future profits gains to be \$350 (SE \$195). The structural estimate for the change in V (\$83) therefore lies within the point estimate's confidence interval.

[48] The willingness to pay measure is designed to capture the present discounted value of profits as perceived by the owner – though it is worth noting that the BDM procedure used to illicit WTP requires owners to have cash on hand to pay for the devices, and this could bias WTP downwards if liquidity constraints limit the amount of money individuals have available (McKenzie and Ubfal, 2020). Our estimate of WTP (\$44.67) is comparable to the empirical estimate of \$350, and the model estimate of \$83, when we factor in the uncertainty in the data.

system, many drivers brought up that trust between owners and drivers could increase under monitoring.

We complement drivers' welfare estimates with SMS survey responses that we collected from drivers six months after the study finished.[49] 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 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 larger amount to their driver in a trust game at endline (Table 5, column 1). This evidence suggests that an improvement in the owner-driver relationship may have counteracted some of the costs drivers' incurred from monitoring. These findings are in line with our model extension on risk aversion, which indicates that when owners cannot accurately infer driver behavior from the noisy signals they receive, they may draw incorrect conclusions and impose excessive punishments that negatively impact driver satisfaction. The introduction of monitoring technologies effectively reduces the frequency of such incidents. This, in turn, has the potential to enhance drivers' trust in the owner and significantly improve their overall satisfaction and happiness.[50]

The welfare implications associated with these monitoring technologies are further complicated by how they interact with the public transit passengers and other road users. One of the motivations for this research initiative was to understand whether these technologies could improve road safety as Kenya's matatu sector is notoriously unsafe. While we explore the implications of GPS technologies on road safety in a companion paper, it is important to highlight that the welfare impacts that we estimate in this paper have the potential to change dramatically if passengers/pedestrian welfare is also considered. Weighing driver welfare relative to consumer welfare is beyond the scope of this research.

[49] 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.

[50] Further integrating the idea of trust in the model and quantifying its importance for welfare is beyond the scope of this paper. One could attempt to model owner and driver utility as a function of deception: there could then be a psychic cost to the driver of deceiving the owner, and the owner suffers from revealed deception. Monitoring would lower the scope for deception and thereby improve welfare.

9 Conclusion

A firm's success rides heavily on the performance of its employees. It is therefore important that firms design contracts and manage their employees in ways that align the employees' incentives with their own. 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 is unclear.

In this paper, we investigate the impact of monitoring devices among small businesses in Nairobi, Kenya. This question is particularly relevant as small firms struggle to expand in low income countries, and information technologies are becoming ubiquitous. To this end, we implement a randomized control trial where we introduce a monitoring device to 255 firms operating in Kenya's transit industry. We design a novel mobile application that provides information to 125 treatment firms regarding: the location of the vehicle, number of kilometers driven, number of hours the ignition was on, and the number of safety violations incurred. We confirm that 70% of owners consult the app weekly. Owners also report that monitoring their drivers has become significantly easier. We use daily surveys from vehicle owners and drivers over six-months to track the impact of reducing asymmetric information on firm outcomes.

Firms use the monitoring device to demand a new bundle of effort and risk from the driver that was previously impossible to incentivize. The driver responds by driving an additional hour per day, and engaging in less off-road driving on bumpy routes that damage the minibus. Vehicle repair costs fall by 45%, and firm profits increase by 13%. These gains more than offset the cost of the device, suggesting that a tracking device like the one we designed for this study would be a worthwhile investment if it were available on the market. We also investigate whether this improved profitability and better management fuel business growth. We find weak statistical evidence that treatment owners have 0.129 more vehicles (10%) on average than control owners after six months.

We do not see the owners changing the contract structure they offer. There is some indication that firms might be reducing the transfer they demand from drivers to compensate them for the higher disutility they incur under the new effort-risk bundle, but the target contracts remains. This suggest that this class of inefficient contracts could remain widespread in this industry where revenue is unobserved– at least until monitoring technologies can reveal the amount of revenue employees earn throughout the day.

To identify the distributional consequences of these technologies we estimate the target contract model via generalized method of moments. Albeit imprecisely estimated, we find that owners' welfare increases by approximately $83 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 the 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 monitoring devices have the potential to help small firms overcome inefficiencies created by moral hazard. These results are particularly relevant for small firms, and policymakers focused on helping firms expand. We know that firms struggle to grow in low income countries for a number of reasons, and this paper identifies another way firms might be able to leverage technology to expand.

References

Akcigit, Ufuk, Harun Alp, and Michael Peters (2021): "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries," *American Economic Review*, Vol. 111, pp. 231–75.

American Bar Association, (ABA) (2018): "How Much Employee Monitoring Is Too Much?," https://www.americanbar.org/news/abanews/publications/youraba/2018/january-2018/how-much-employee-monitoring-is-too-much-/.

Anderson, C. Leigh, Maya Dietz, Andrew Gordon, and Marieka Klawitter (2004): "Discount Rates in Vietnam," *Economic Development and Cultural Change*, Vol. 52, pp. 873–887.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments*," *The Quarterly Journal of Economics*, Vol. 132, pp. 1553–1592.

Atkin, David, Amit K. Khandelwal, and Adam Osman (2017): "Exporting and Firm Performance: Evidence from a Randomized Experiment*," *The Quarterly Journal of Economics*, Vol. 132, p. 551.

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.

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F Katz, Benjamin A Olken, and Anja Sautmann (2020): "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics," Working Paper 26993, National Bureau of Economic Research.

Bassi, Vittorio, Raffaela Muoio, Tommaso Porzio, Ritwika Sen, and Esau Tugume (2022): "Achieving Scale Collectively," *Econometrica*, Vol. 90, pp. 2937–2978.

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

Berg, Joyce, John Dickhaut, and Kevin McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, Vol. 10, pp. 122 – 142.

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.

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.

Brooks, Wyatt, Kevin Donovan, and Terence R. Johnson (2018): "Mentors or Teachers? Microenterprise Training in Kenya," *American Economic Journal: Applied Economics*, Vol. 10, pp. 196–221.

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.

Caria, Stefano A. and Paolo Falco (2022): "Skeptical Employers: Experimental Evidence on Biased Beliefs Constraining Firm Growth," *The Review of Economics and Statistics*, pp. 1–45.

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.

Dickens, William T., Lawrence F. Katz, Kevin Lang, and Lawrence H. Summers (1989): "Employee Crime and the Monitoring Puzzle," *Journal of Labor Economics*, Vol. 7, pp. 331–347.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar (2014): "Keeping It Simple: Financial Literacy and Rules of Thumb," *American Economic Journal: Applied Economics*, Vol. 6, pp. 1–31.

Foster, Andrew D. and Mark R. Rosenzweig (1995): “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture,” *Journal of Political Economy*, Vol. 103, pp. 1176–1209.

Gertler, Paul, Marcelo Olivares, Raimundo Undurraga, and Pablo Celhay (2023): “Monitoring, Organizational Culture, and Procurement Efficiency,” *Working Paper*.

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.

Gollin, Douglas (2002): “Getting Income Shares Right,” *Journal of Political Economy*, Vol. 110, pp. 458–474.

Gosnell, Greer K., John A. List, and Robert D. Metcalfe (2020): “The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains,” *Journal of Political Economy*, Vol. 128, pp. 1195–1233.

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.

Hanna, Rema, Gabriel Kreindler, and Benjamin A. Olken (2017): “Citywide effects of high-occupancy vehicle restrictions: Evidence from “three-in-one” in Jakarta,” *Science*, Vol. 357, pp. 89–93.

Hardy, Morgan and Jamie McCasland (2020): “Are Small Firms Labor Constrained? Experimental Evidence From Ghana,” *Working Paper*.

Hsieh, Chang-Tai and Benjamin A. Olken (2014): “The Missing ”Missing Middle”,” *Journal of Economic Perspectives*, Vol. 28, pp. 89–108.

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.

Jayachandran, Seema (2020): "Microentrepreneurship in Developing Countries," *Working Paper*.

Kalan, Jonathan (2013): "The Technology Modernising Kenya's Matatus."

Karlan, Dean and Martin Valdivia (2011): "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions," *The Review of Economics and Statistics*, Vol. 93, pp. 510–527.

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

Kreindler, Gabriel (2020): "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," *Working Paper*.

Lane, Gregory, David Schönholzer, and Erin M. Kelley (2022): "Information and Strategy in Lemon Markets: Improving Safety in Informal Transit," *Working Paper*.

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

Liu, Meng, Erik Brynjolfsson, and Jason Dowlatabadi (2021): "Do Digital Platforms Reduce Moral Hazard? The Case of Uber and Taxis," *Management Science*, Vol. Forthcoming.

Lucas, Jr., Robert E. (1978): "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, Vol. 9, pp. pp. 508–523.

Macharia, WM, EK Njeru, F Muli-Musiime, and V Nantulya (2009): "Severe Road Traffic Injuries in Kenya, Quality of Care and Access," *African Health Sciences*, Vol. 9.

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 (2017): "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition," *American Economic Review*, Vol. 107, pp. 2278–2307.

McKenzie, David and Anna Luisa Paffhausen (2019): "Small Firm Death in Developing Countries," *The Review of Economics and Statistics*, Vol. 101, pp. 645–657.

McKenzie, David and Diego Ubfal (2020): "Using BDM and TIOLI to measure the demand for business training in Jamaica."

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

McKenzie, David, Christopher Woodruff, Kjetil Bjorvatn, Miriam Bruhn, Jing Cai, Juanita Gonzalez-Uribe, Simon Quinn, Tetsushi Sonobe, and Martin Valvidia (2021): "Training Entrepreneurs: Issue 2 — VoxDev," https://voxdev.org/lits/training-entrepreneurs.

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

Odero, Wilson, Meleckidzedeck Khayesi, and P. M. Heda (2003): "Road Traffic Injuries in Kenya: Magnitude, Causes and Status of Intervention," *Injury Control and Safety Promotion*, Vol. 10, pp. 53–61, PMID: 12772486.

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

Shahe Emran, M., A. K. M. Mahbub Morshed, and Joseph E. Stiglitz (2021): "Microfinance and Missing Markets," *Canadian Journal of Economics/Revue canadienne d'économique*, Vol. 54, pp. 34–67.

Sprenger, Charles (2015): "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, Vol. 123, pp. 1456–1499.

Startz, Meredith, Gregory Fischer, Dean Karlan, and Diego Santa Maria (2023): "Business Training and Spatial Competition in Urban Uganda," *Working Paper*.

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

Tsivanidis, Nick (2019): “Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá’s TransMilenio,” *Working Paper*.

U.S. Department of Transportation (2016): “National Transportation Statistics,”Technical report, Bureau of Transportation Statistics.

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.

Van Lieshout, S and P Mehtha (2017): “The Next 15 Million: Start and Improve Your Business Global Tracer Study 2011-2015”,”Technical report, International Labour Organization, Geneva.

West, Darell M. (2021): “How Employers use Technology to Surveil Employees,” *Brookings Techtank*.

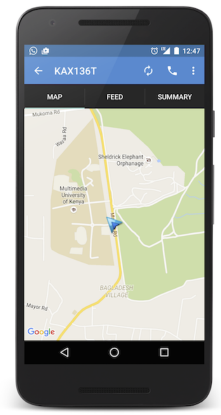
WHO (2013): “Global status report on road safety 2013: supporting a decade of action,”Technical report, World Health Organization.

World Bank (2021): “Small and Medium Enterprises (SMEs) Finance: Improving SMEs’ Access to Finance and Finding Innovative Solutions to Unlock Sources of Capital,” June.

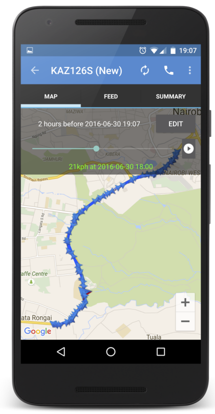
Figures

Figure 1: Mobile app "SmartMatatu"

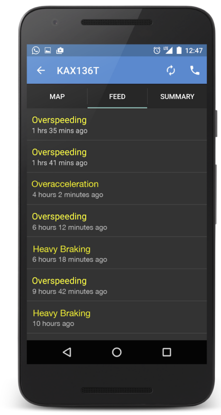
(a) Map Viewer



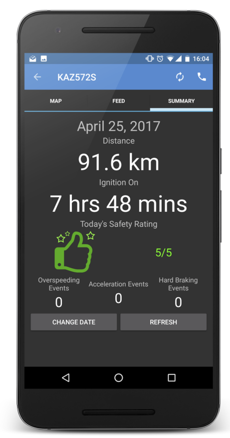
(b) Historical Map Viewer



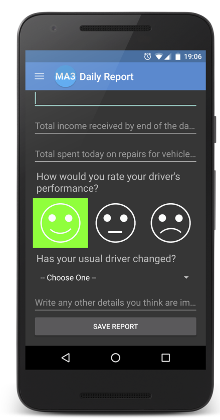
(c) Safety Feed



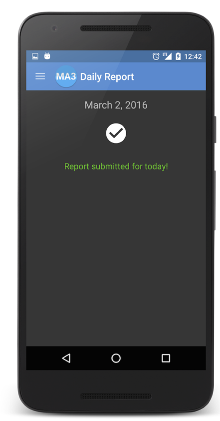
(d) Effort Summary



(e) Report Submit



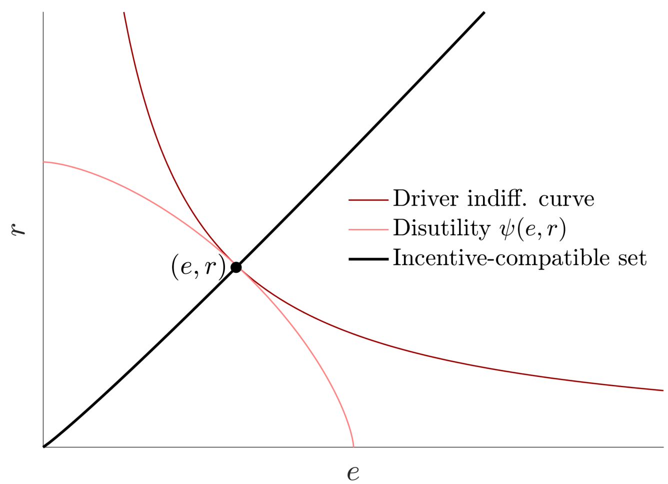
(f) Report Complete



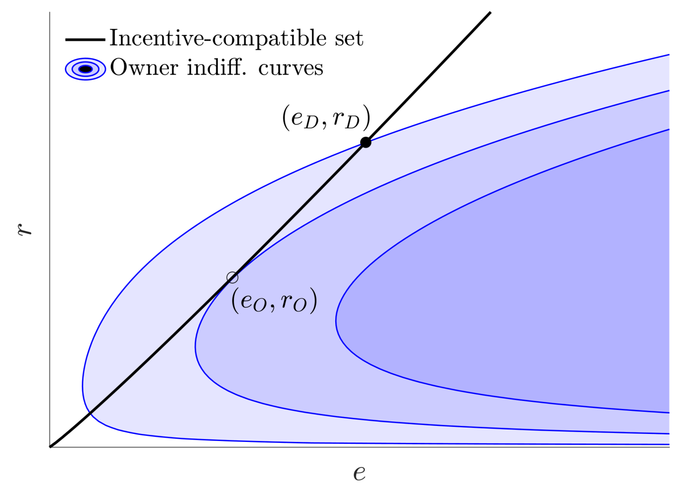
Notes: This figure presents the Android mobile app "SmartMatatu" developed by Echo Mobile in collaboration with matatu owners. Panels A and B: map viewer of real-time matatu location with historical playback of past locations over several hours for a given day. Panel C: Safety feed with speeding, acceleration, and hard braking alerts. Panel D: Daily effort 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

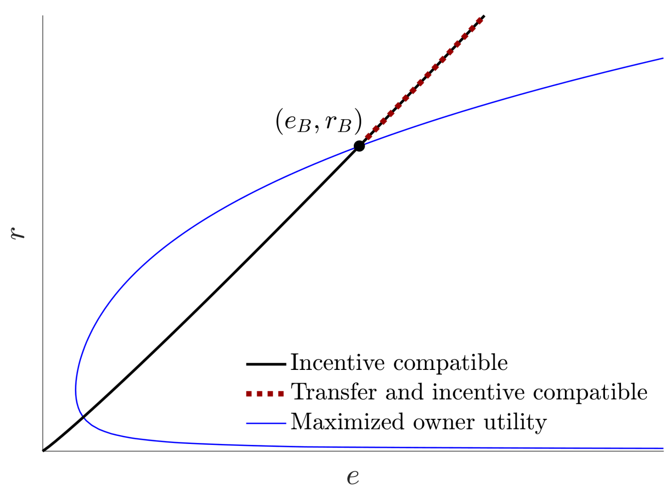
(a) Driver utility in effort-risk (e, r) space.



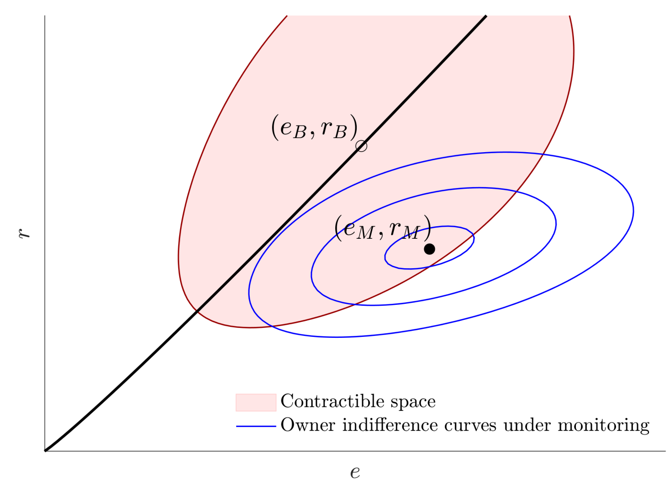
(b) Owner prefers lower risk: $(e_D, r_D) > (e_O, r_O)$.



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

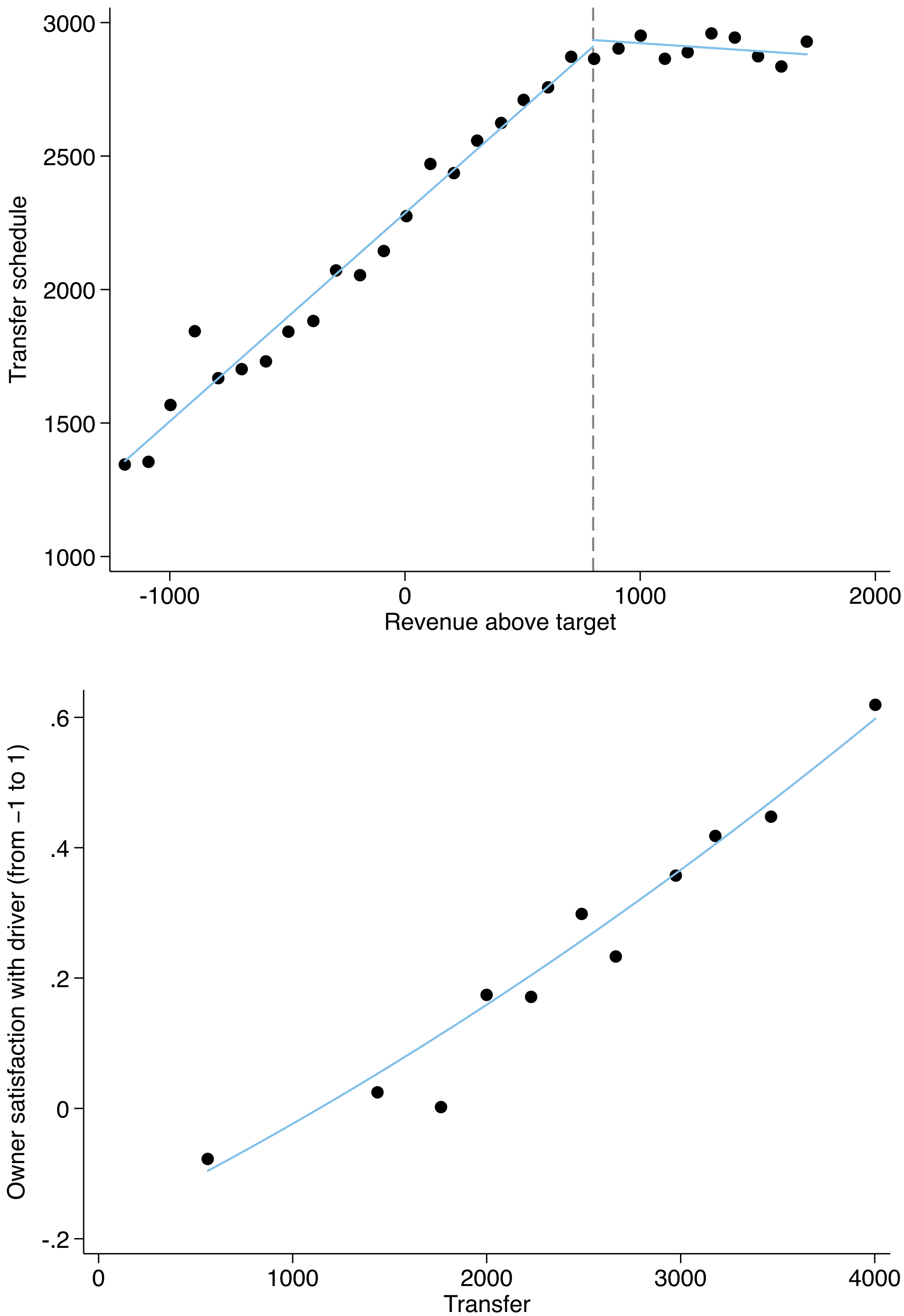


(d) Monitoring shifts effort/risk to (e_M, r_M).



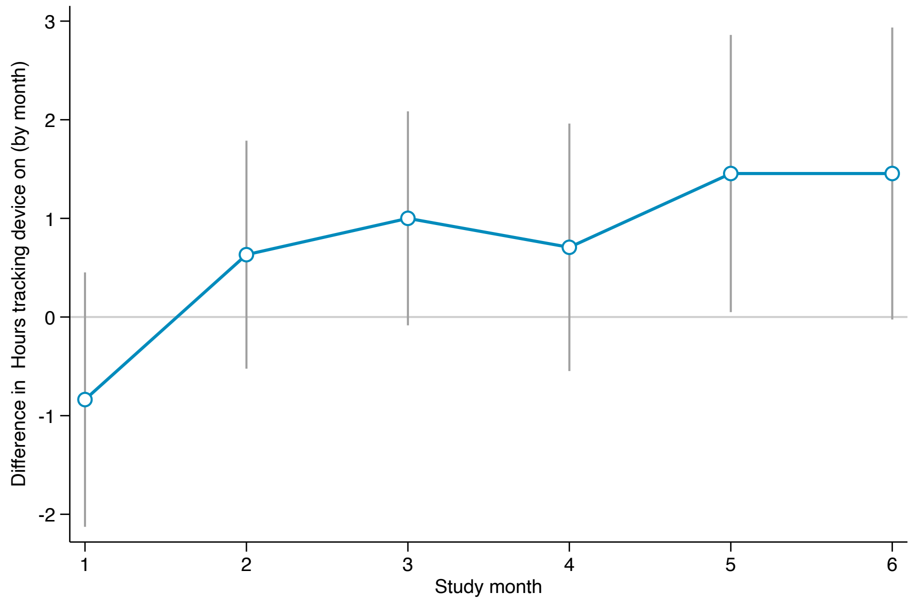
Notes: Panel A: Without monitoring, the owner can only induce effort-risk bundles on the incentive compatible set. Panel B: The owner's preferred driving choices (e_O, r_O) exhibit lower risk *and* effort than the driver's (e_D, r_D) on the incentive-compatible set (see Assumption 2). Panel C: The baseline 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

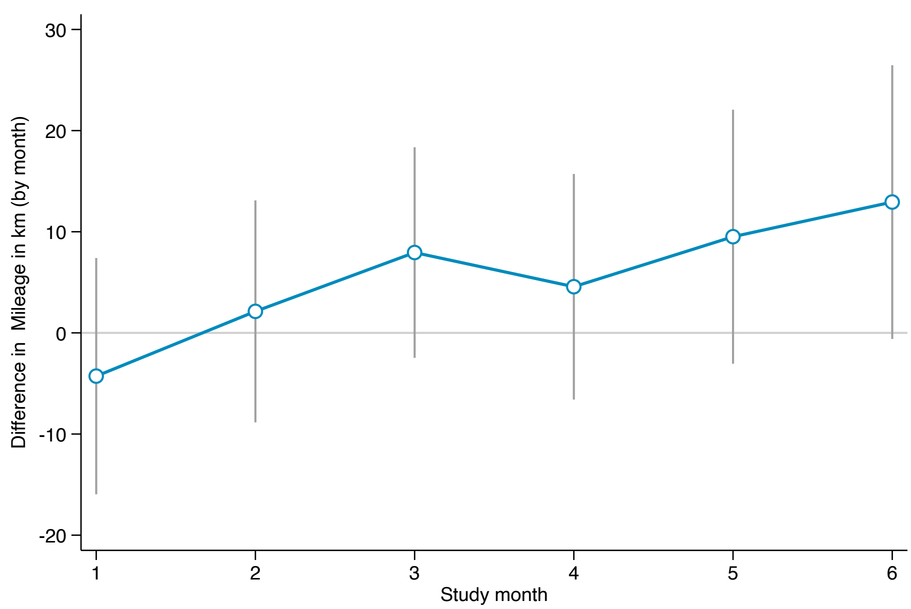


Notes: Top panel: The empirical transfer schedule as a function of the amount of revenue earned (above the target). This empirical transfer schedule closely resembles the shape $t(y) = \min\{y, T\}$ as in the Lemma. The slope extends beyond the target because of subsistence income, which we include in the structural estimation (see text). Bottom panel: Owner satisfaction with driver as a function of the transfer. Owner satisfaction rises substantially with the transfer, as suggested by $p(t) = 1 - \frac{T-t}{\delta U - \bar{u}}$.

Figure 4: Treatment effects on effort



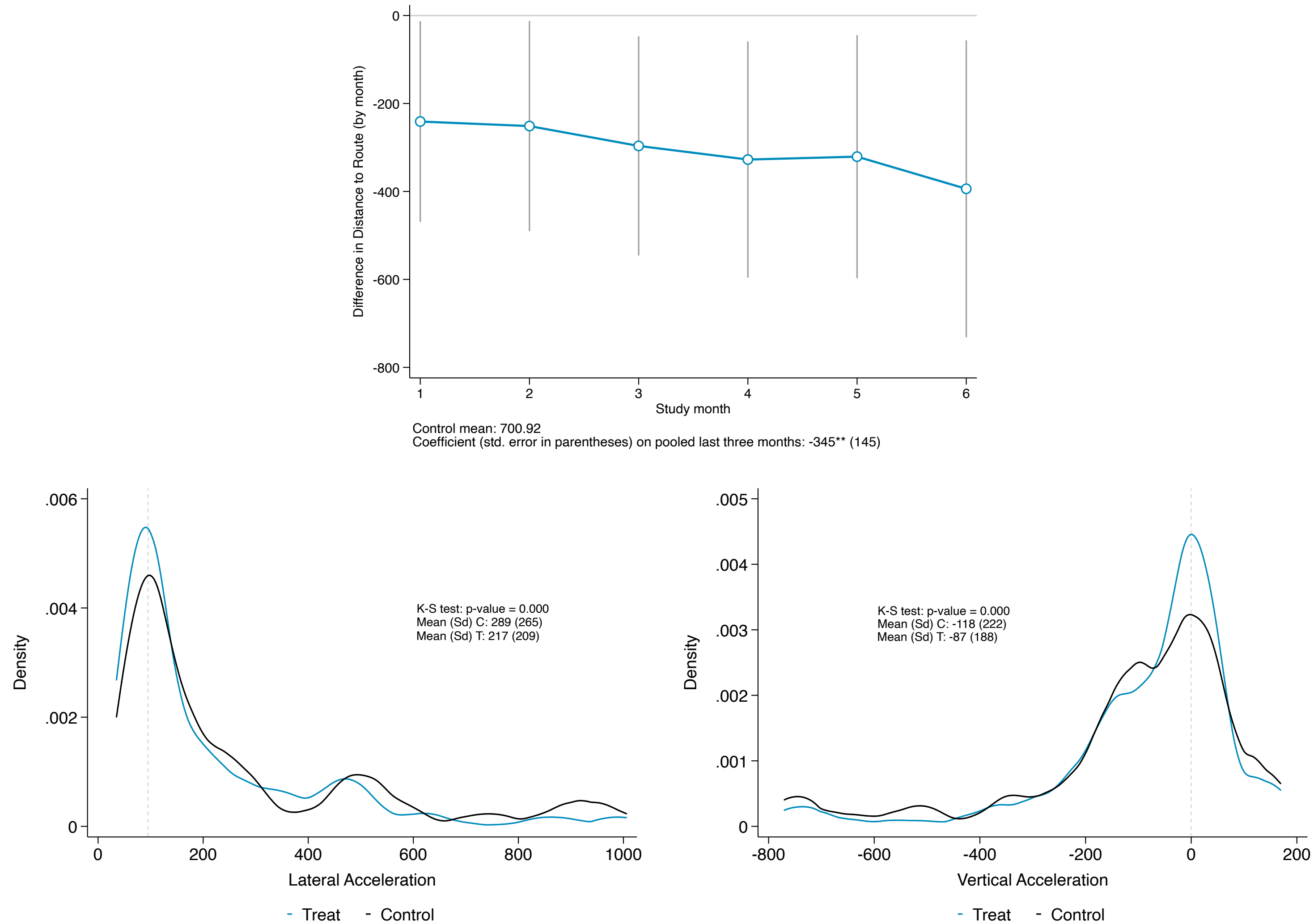
Control mean: 14.79
Coefficient (std. error in parentheses) on pooled last three months: 1.11 (0.64)



Control mean: 96.64
Coefficient (std. error in parentheses) on pooled last three months: 8.32 (5.79)

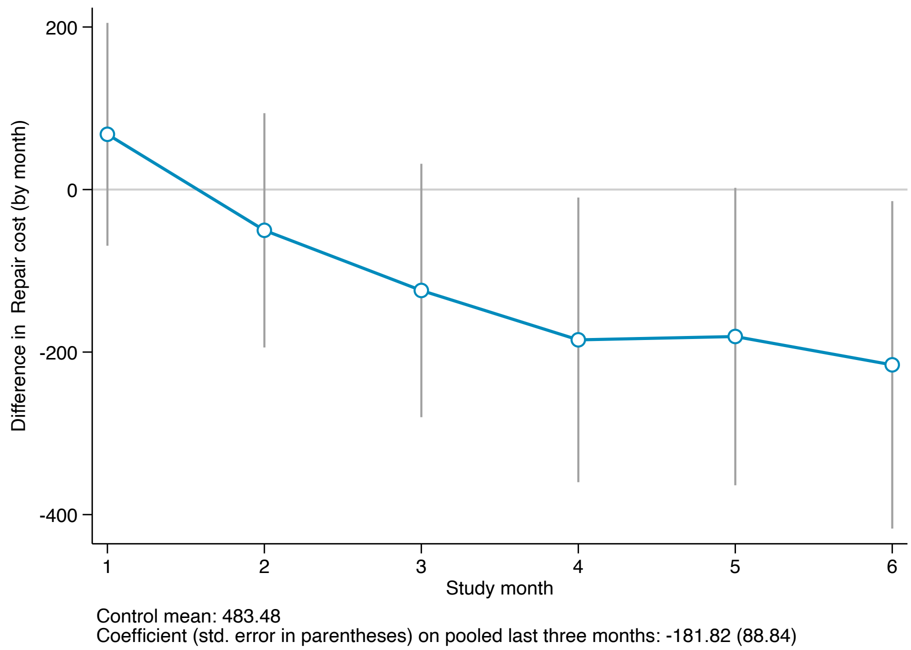
Notes: OLS estimates according to Equation 2. Top panel: Treatment effect by month on hours tracking device on, which corresponds to working hours of driver. Bottom panel: Treatment effect by month on daily mileage captured by tracking device. Standard errors for 95% confidence intervals clustered at the matatu level. In each graph, we present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure 5: Treatment effects on risk taking



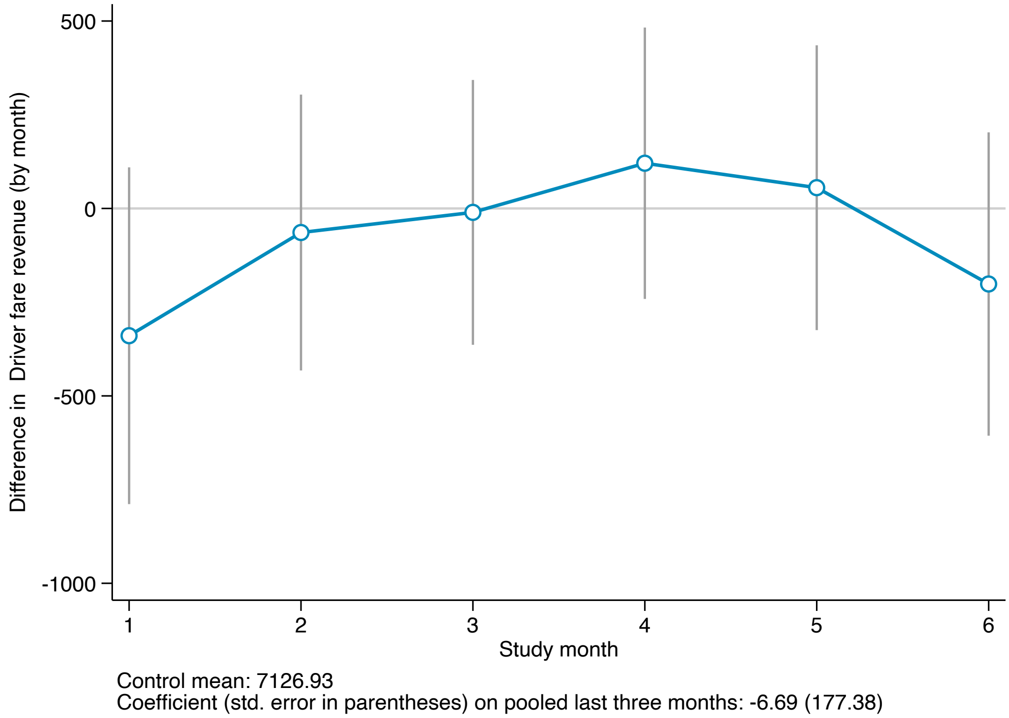
Notes: Top panel: OLS estimates according to Equation 2. Treatment effect by month on distance to licensed route in meters captured by tracking device. Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2). Bottom panel: treatment (blue) and control (black) distributions of lateral (left) and vertical (right) acceleration. We present Kolmogorov-Smirnov tests of equality of these distributions across treatment and control.

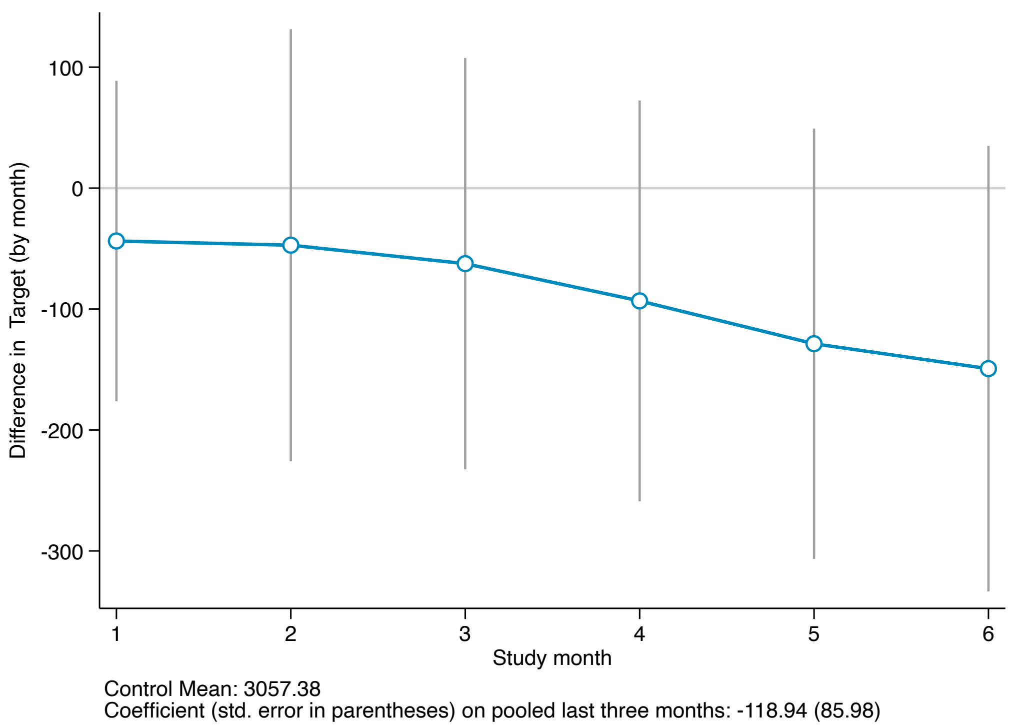
Figure 6: Treatment effects on costs



Notes: OLS estimates according to Equation 2. Treatment effect by month on costs, defined as the repair costs reported by the owners. Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

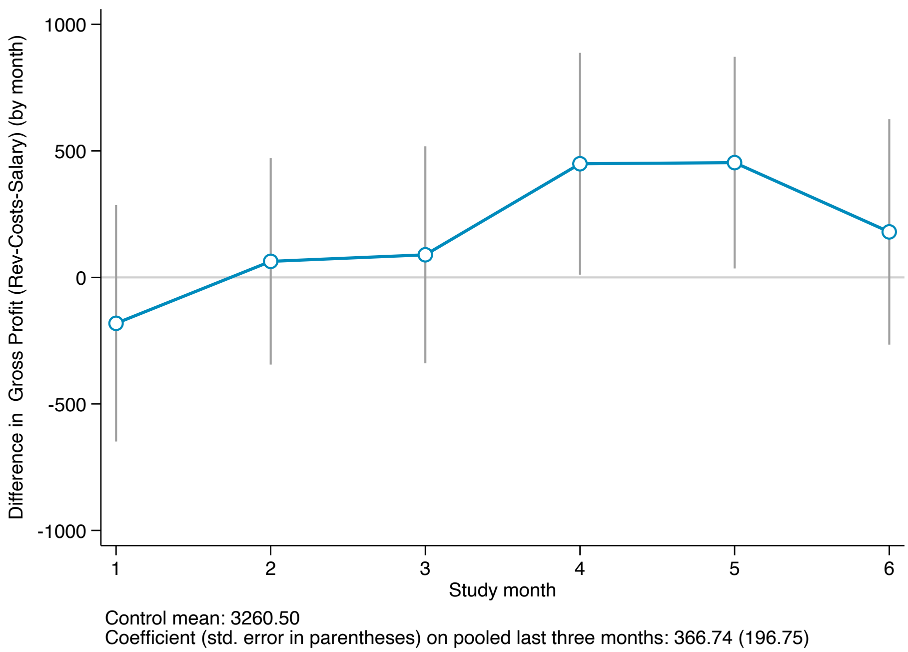
Figure 7: Effect of monitoring on revenue and target





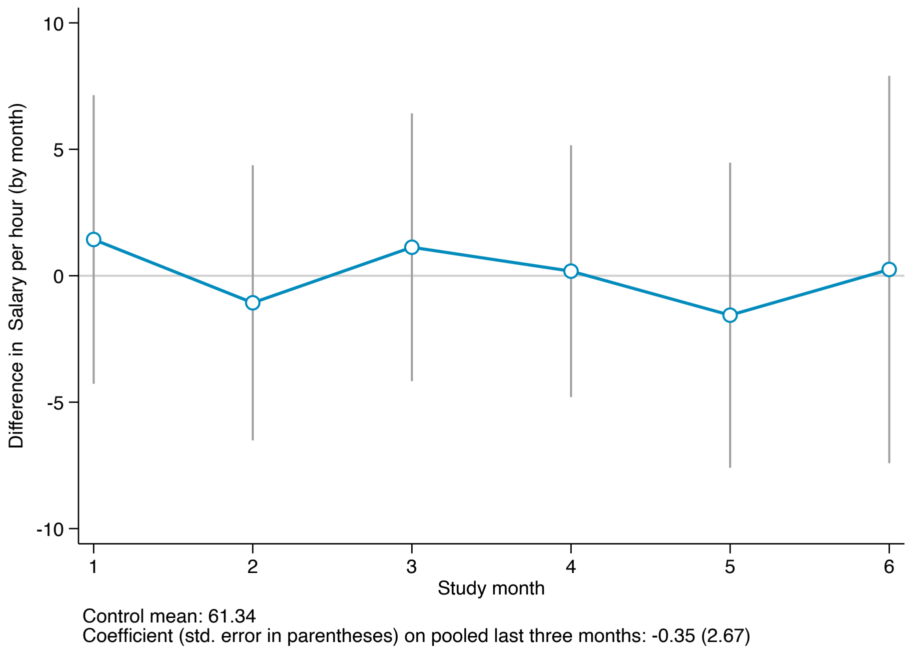
Notes: OLS estimates according to Equation 2. Top panel: Treatment effects by month on daily revenue reported by the driver. Bottom panel: Treatment effects by month on the target amount the owner assigns to their driver at the beginning of the day. Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure 8: Treatment effects on profits



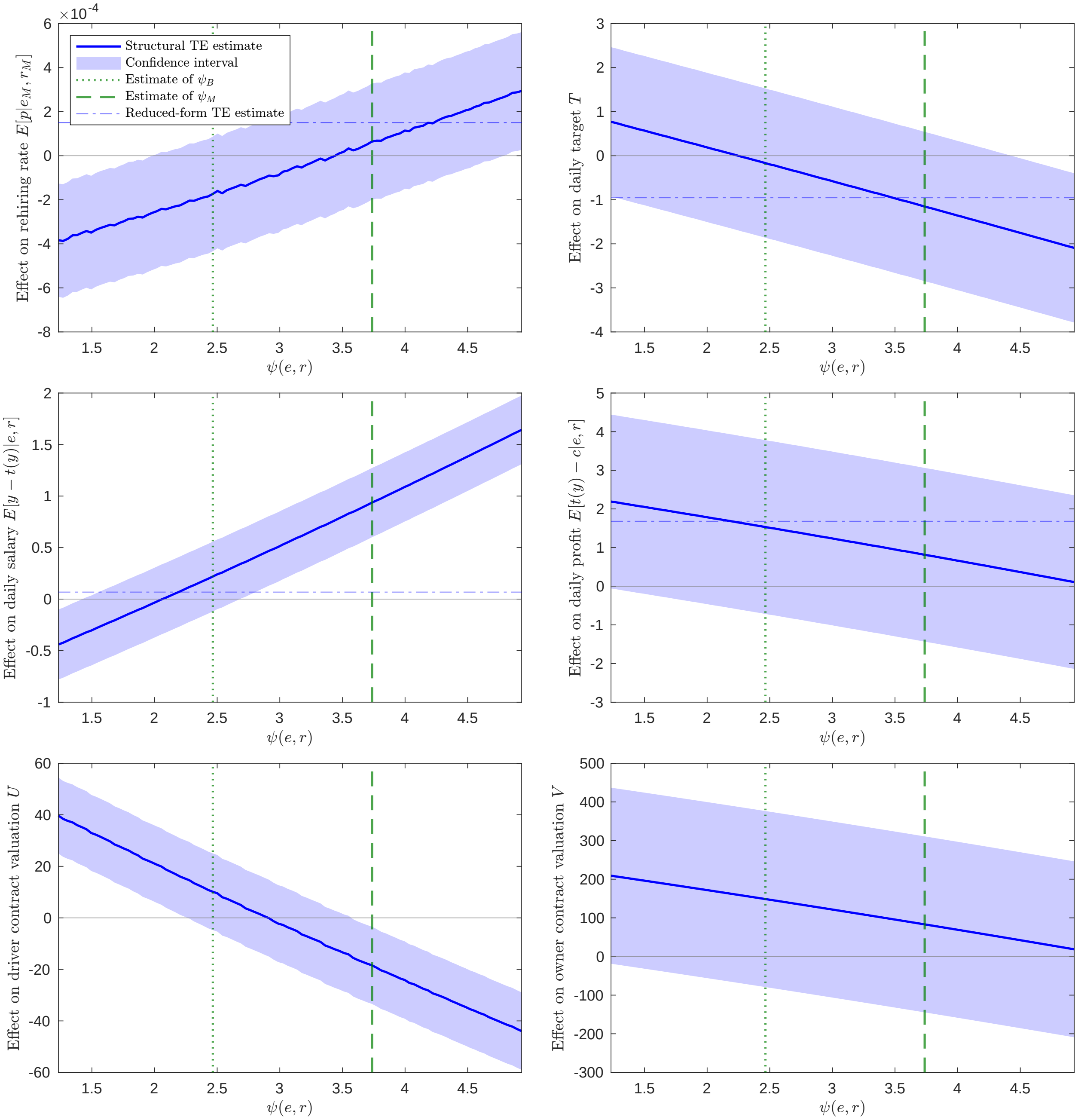
Notes: OLS estimates according to Equation 2. Treatment effect by month on gross profit, defined as revenue minus repair costs minus driver residual claim (salary). Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure 9: Treatment effects on salary per hour



Notes: OLS estimates according to Equation 2. Treatment effect by month on driver salary per hour. Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure 10: Structural treatment effect estimates due to increase in driver disutility



Notes: Figures plot the structural treatment effect estimates for driver disutility $\psi(e, r)$ ranging from 50% to 200% of the baseline value. The two panels on the top (rehiring probability and target) are the targeted moments, whereas the other four moments are untargeted. Dotted line shows estimated baseline disutility ψ_B, and dashed line shows estimated disutility under monitoring ψ_M. The dash-dotted horizontal line shows reduced-form treatment effect estimates in the cross-section, where applicable (these may differ from the estimates in the experimental estimates due to the lack of controls, fixed effects, and weighting by number of bus-days). The other two parameters ($\bar{u}$ and h) are fixed at the values estimated in the control sample.

Tables

Table 1: Summary statistics for 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	
Joint Test							0.332

Notes: This table presents summary statistics for the owners, drivers, and matatus in our sample. We report mean, standard deviation, min and max for the full sample (column 1,2,3,4), means for the treatment and control groups (column 5,6) and the p-value of the t-test comparing means in treatment and control groups. (column 7). Baseline target, baseline transfer, baseline revenue, baseline residual revenue are in 100s of Kenyan Shillings (KES, approximately $1). Years of education is 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. Raven's score represents the respondent's score on a cognitive assessment. Data from baseline survey.

Table 2: Treatment effects on reported knowledge and monitoring behavior

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

Notes: This table shows the impact of treatment on owners' knowledge and monitoring practices (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": a binary (yes/no) for whether the owner reports knowing the approximate number of kilometers a driver drove on a given day. "Know off-route": a binary (yes/no) for whether the owner knows when the driver is off the licensed route. "Know revenue": a binary (yes/no) for whether 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 endline survey. These additional questions were 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). Robust standard errors.

Table 3: Treatment effects on effort, costs, revenue, target, profits, salary

	(1) Device on (hours)	(2) Mileage (kilometers)	(3) Repair costs	(4) Repair costs (large)	(5) Revenue	(6) Target	(7) Met target	(8) Gross profit	(9) Salary per hour
Treatment × Month 1	-0.84	-4.28	68.0	0.033	-339.5	-43.7	-0.100	-181.6	1.43
	(0.66)	(5.96)	(70.0)	(0.030)	(229.2)	(67.6)	(0.045)	(238.3)	(2.91)
Treatment × Month 2	0.63	2.13	-50.2	-0.019	-64.2	-47.2	-0.012	63.3	-1.07
	(0.59)	(5.60)	(73.5)	(0.030)	(187.7)	(91.1)	(0.046)	(208.1)	(2.78)
Treatment × Month 3	1.00	7.94	-124.2	-0.030	-10.5	-62.5	0.071	89.2	1.13
	(0.55)	(5.31)	(79.5)	(0.032)	(180.2)	(86.8)	(0.048)	(218.8)	(2.70)
Treatment × Month 4	0.71	4.56	-185.0	-0.047	120.6	-93.3	0.13	449.0	0.18
	(0.64)	(5.69)	(89.4)	(0.033)	(184.7)	(84.6)	(0.051)	(223.6)	(2.54)
Treatment × Month 5	1.45	9.51	-180.9	-0.064	55.4	-128.7	0.080	453.5	-1.56
	(0.72)	(6.41)	(93.4)	(0.033)	(193.8)	(90.8)	(0.054)	(213.3)	(3.08)
Treatment × Month 6	1.45	12.9	-215.7	-0.077	-201.6	-149.3	0.054	179.8	0.25
	(0.76)	(6.90)	(102.8)	(0.038)	(206.4)	(94.0)	(0.054)	(227.3)	(3.91)
Control Mean of DV	14.8	96.6	483.5	0.17	7126.9	3057.4	0.43	3260.5	61.3
Joint Test	0.03	0.09	0.01	0.03	0.15	0.78	0.00	0.02	0.83
Controls	X	X	X	X	X	X	X	X	X
Day FE	X	X	X	X	X	X	X	X	X
Route FE	X	X	X	X	X	X	X	X	X
Matatu-Day N	45,654	45,654	15,881	15,881	22,436	15,888	15,888	10,406	22,426

Notes: This table presents treatment effects for all the experimental results (OLS regressions as in Equation 2). "Device on (hours)": number of hours the tracking device reported the ignition to be on. "Mileage (kilometers)": number of kilometers the tracking device reported the bus on the road. "Repair costs": owner-reported daily repair costs. "Large repair costs": owner-reported daily repair costs that exceed 10 USD. "Revenue": driver-reported daily revenue. "Target": daily revenue target set by owner. "Met target": whether the driver met the target. "Gross profit": Revenue minus repair costs minus driver residual claim (salary). "Salary": driver-reported residual claim (salary). 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, driver SMS reports, and aggregated tracking device data. We report a joint test of all six monthly treatment coefficients. Standard errors clustered at the owner/driver/matatu level.

Table 4: 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: This table shows the impact of treatment on business investment (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). "Number vehicles": the number of vehicles the owner owns at endline. "New interior": whether a major investment into interior of vehicle was made. Data from endline survey. Robust standard errors.

Table 5: Treatment effects on owner's perceptions of their drivers' performance

	(1) Trust amount	(2) More honest	(3) Performance rating	(4) Better driving
Treatment	33.80	0.71	0.11	0.63
	(15.12)	(0.05)	(0.17)	(0.06)
Control Mean of DV	151.61	0.04	7.21	0.04
Controls	X	X	X	X
Route FE	X	X	X	X
Matatu N	244	190	246	190

Notes: This table shows how treatment affects owners' perceptions of their drivers' performance (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": owner's perception of whether driver's honesty has changed since baseline, is 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). "Better driving": the owner's judgement of overall driver performance at endline, worse (-1), about the same (0), or better (1). Data from endline survey. Questions about honesty and better driving were added after a quarter of endlines were already completed, hence only up to 190 out of 255 observations (balanced across treatment and control). Robust standard errors.

Table 6: Model estimation under baseline contract

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on control group
Discount factor δ	0.99	

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Baseline driver disutility $\psi(e_B, r_B)$	2.47	Driving disutility of \$2.47
	(0.70)	
Firing cost h	263	Lost profit of firing of \$263
	(19)	(about 11 days of profit)
Driver outside option $\bar{u}$	1.57	Similar to unskilled daily wage with
	(0.37)	subsistence (\$1.57 + \$7 = \$8.57)

Panel C: Reduced form, structural, and matched moments			
Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Firing probability $E[p\|e, r]$	0.007	0.007	-0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	507.1	-0.7
	(12.9)	(70.4)	(71.6)
Target T	30.1	30.1	0.0
	(0.4)	(0.7)	(0.8)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	9.1	9.9	-0.8
	(0.2)	(0.5)	(0.5)
Owner profit $E[t\|e, r] - E[c\|e, r]$	24.3	23.5	0.8
	(0.7)	(0.5)	(0.9)
Owner contract value V	—	2,177	—
	—	(10)	—
Welfare $U + V$	—	2,684	—
	—	(71)	—

Notes: Generalized method of moments (GMM) estimation of driver disutility, firing cost, and outside option. Sample: control group. Targeted moments: Separation probability, driver contract value, and target. "Reduced form" as observed in the sample. "Structural" are the corresponding estimated model predictions. The difference is between reduced form and structural moments. Standard errors of parameters based on estimated asymptotic variance and of structural moments via the Delta Method.

Table 7: Reduced form versus structural treatment estimation

Panel A: Assumptions

Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e,r)$	—	Normal distribution on treated group
Discount factor δ	0.99	
Outside option $\bar{u}$	1.57	Estimated in control group
Firing cost h	263	Estimated in control group

Panel B: GMM parameter estimates

Input	Value	Interpretation
Disutility with monitoring $\psi(e_M, r_M)$	3.74	Increase of $1.27 (52%)
	(1.94)	

Panel C: Reduced form, structural, and matched treatment effects

Treatment effect	Red. form (Δ)	Structural (Δ)	Difference (Δ)
Targeted moments:			
Firing probability $E[p\|e,r]$	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Target T	-1.0	-1.1	0.2
	(0.6)	(1.1)	(1.3)
Untargeted moments:			
Driver salary $E[y\|e,r] - E[t\|e,r]$	0.1	0.9	-0.9
	(0.2)	(0.7)	(0.7)
Owner profit $E[t\|e,r] - E[c\|e,r]$	1.7	0.8	0.9
	(1.0)	(1.3)	(1.6)
Driver contract value U	—	-19.6	—
	—	(22.4)	—
Owner contract value V	—	83.1	—
	—	(124.7)	—
Welfare $U + V$	—	63.5	—
	—	(129.0)	—

Notes: Driver disutility under monitoring estimated via GMM. Sample: treatment group. Targeted moments: Separation probability and target. Untargeted moments: driver contract value, driver salary, owner profit. "Reduced form (Δ)" are the difference between the treatment group and the control group in the data. "Structural (Δ)" are the corresponding difference between estimated model predictions of the treatment and control groups. "Difference (Δ)" is the difference between between reduced form and structural moment differences. Standard errors via the bootstrap.

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

Erin Kelley, Gregory Lane, and David Schönholzer

Contents

A Ethics Statement

This paper is focused on Kenya's public transit industry, a sector dominated by small-scale entrepreneurs that typically own 1-2 minibuses ("matatus") which they rent out to drivers on a daily basis. Over the past 10 years new technologies have entered the market and changed the way minibus owners manage their businesses. Most notably are the arrival of GPS technologies that help matatu owners track their vehicles. When we launched the study in 2014 these technologies were still relatively new. Private businesses were encouraging owners to purchase tracking devices, and designing applications that customers could download to monitor matatu's progress (Kalan, 2013). Finally, a number of banks were requiring that matatu owners install GPS trackers in their minibuses before approving a loan for a new bus.

While these new technologies were spreading rapidly, there were no active studies investigating their impact on the transportation industry as a whole. As a result, we felt that it was important to document their impact on matatu owners, drivers and passengers/commuters. The latter are an important group to consider as the minibus industry is notoriously unsafe and road traffic accidents are becoming the leading causing of death among 18-25-year olds in low-income countries (WHO, 2013). We explore these dynamics on road safety in a companion paper.

A priori there were two major concerns about the impact of the device on matatu drivers. First, there was a concern that drivers could lose their jobs, or some of their income as a result of the information that we provided to matatu owners. After more than a year of piloting in the field we collected sufficient evidence to suggest these outcomes were unlikely. Extensive conversations with owners and drivers highlighted that drivers hold significant market power because finding reliable drivers that owners can trust is not easy. In most cases of owner-driver separations, drivers reported leaving voluntarily for a better paying job from another owner. Furthermore, of the relationships that were terminated by the owner, drivers reported being able to find a similarly paying job quickly and without difficulty. Conversely, owners reported significant difficulties in finding drivers to operate their matatus. Therefore, we expected that owners would not financially punish their drivers in response to this information.

Despite these assurances, there remained some risk to drivers along this dimension. To minimize the chance that matatu drivers were negatively affected by the intervention we set up a hotline that drivers could use to contact us at any time. We informed them at baseline that they should use this number to notify us if the owners threatened to act in a harmful manner. We hired a well-established matatu driver's advocate to attend the baseline surveys and explain the value of this resource. This helped build the requisite trust with drivers. We also made monthly phone calls to each driver to check-in on business operations. These precautions were successful. We did not record any instance of drivers using the hotline or reporting abusive behavior. Moreover, driver income was not affected by our intervention (and if anything, it increased ever so slightly).

The second concern we had is that owners would use the GPS technology to mandate a new level of effort and risk that drivers would find burdensome and make them worse off. A priori, however, the effect on this outcome was ambiguous. When owners have access to a GPS technology they can monitor dimensions of driver behavior that were previously unobserved. This broadens the contracting space and could result in the owner choosing an effort/risk/target profile that is more appealing to the driver, leaving the driver better off.

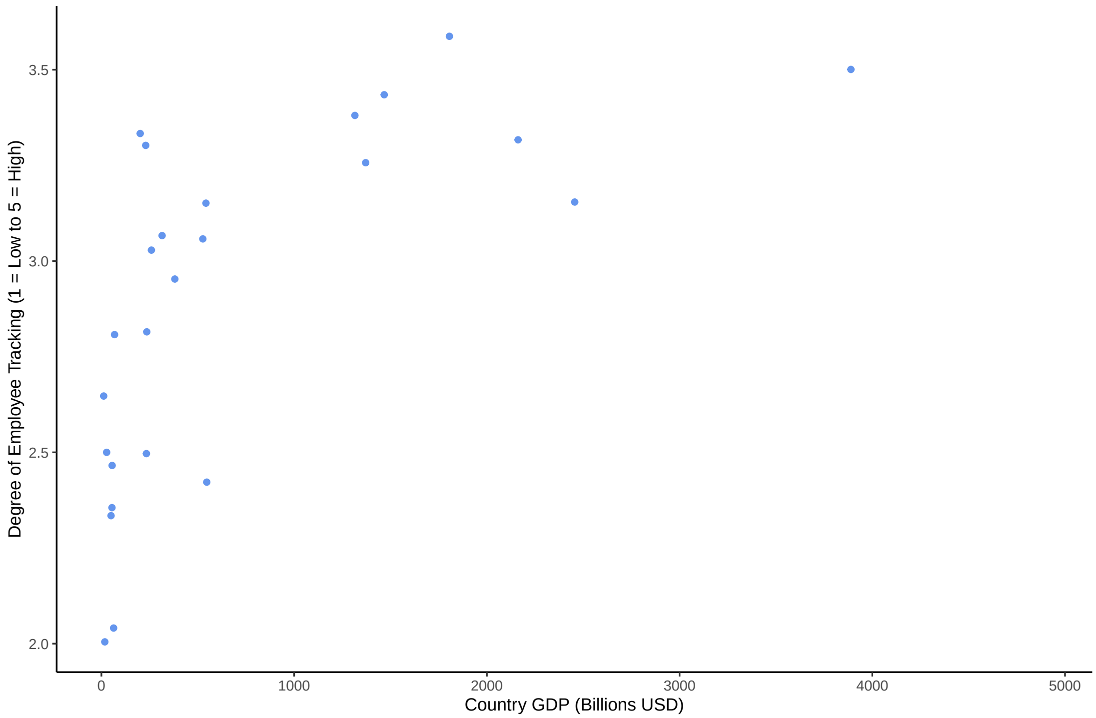
While we find that owners end up offering a welfare reducing contract to drivers in this case, there is an important caveat to this finding. These welfare estimates do not account for the intangible relationship between owners and drivers. The relationship between owners and drivers in is notoriously fraught with mistrust. Matatu drivers often complain that owners second-guess their reports, and refuse to give them the benefit of the doubt when things go awry. Drivers were initially the ones who communicated to the research team that monitoring devices could make these interactions much easier. Other proponents of monitoring technologies also suggest they can foster better working relationships by increasing employers' trust in their employees (Pierce, Snow, and McAfee, 2015). We have some suggestive evidence of this in our data. In a qualitative survey we conducted six months after the experiment concluded, we find that 65% of drivers said the tracking device made their job easier (26% said nothing changed). We also see that owners share an additional 30 KES with drivers in a trust game that we played at endline – a 30% increase. This suggests that the effects of new technologies on worker well-being are more nuanced than what our welfare estimates capture.

The welfare implications associated with these monitoring technologies are further complicated by how they interact with the consumers of public transit and other road users. One of the motivations for this research initiative was to understand whether these technologies could improve road safety. Kenya's matatu sector is notoriously unsafe: drivers often over-accelerate, speed, stop suddenly, and turn sharply in order to collect more passengers. Matatus account for 11% of registered vehicles but 70.2% of passenger casualties (Macharia et al., 2009). Buses in the US on the other hand account for 1% of registered vehicles and

0.4% of casualties (U.S. Department of Transportation, 2016). While we explore the implications of GPS technologies on road safety in a companion paper, it's important to highlight that the welfare impacts that we calculate in this paper have the potential to change dramatically if passengers/pedestrian welfare is also considered. Weighing driver welfare relative to consumer welfare is beyond the scope of this research.

B Appendix Figures

Figure A.1: Employee Tracking and GDP

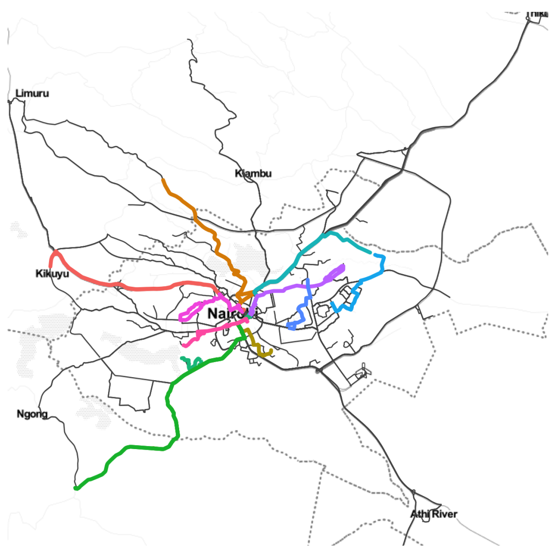


Notes: This figure presents the relationship between the degree of performance tracking by firms in a particular country according to the World Management Survey, and that country's GDP according to the World Bank's Data Catalog for GDP. We see a positive correlation: firms are more likely to use meaningful metrics to track employee performance as GDP improves. The data only include countries for whom the World Management Survey tracks firm outcomes.

Figure A.2: Metropolitan Nairobi matatu route maps

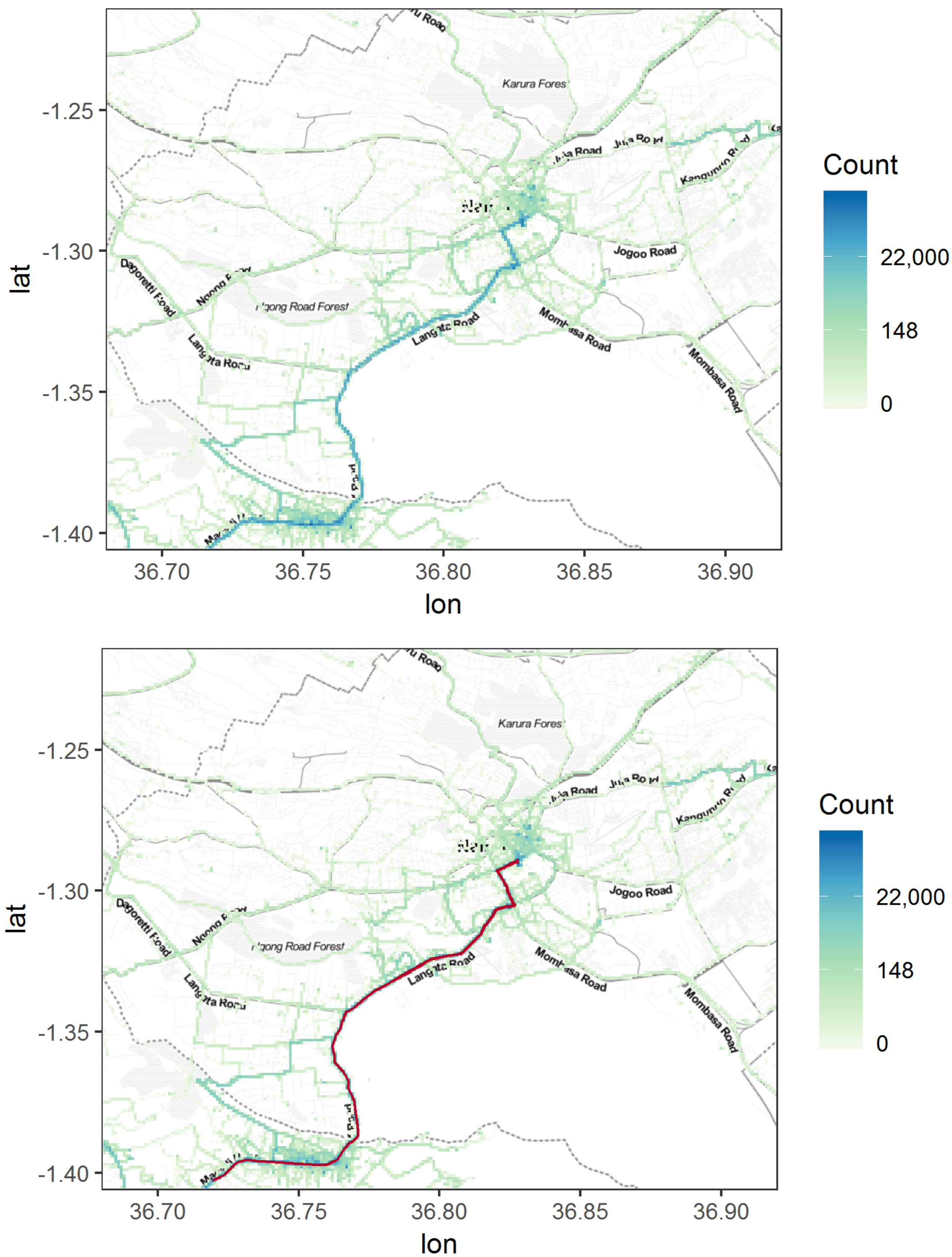
(a) Designated bus routes in Nairobi (black)

(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 routes.

Figure A.3: Device location

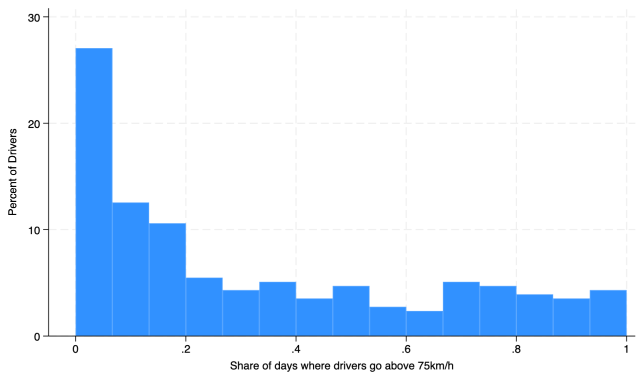


Notes: These maps use data from the trackers that were installed in vehicles licensed to operate on Route 126 (Ongata-Rongai line). We count the number of times that vehicles passed through particular geographic cell on the map. A deeper shade of blue demonstrates that more vehicles passed through that 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.

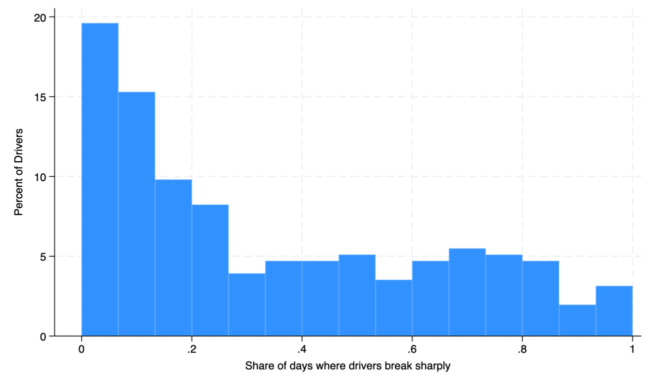
Figure A.4: Example of vehicles used in study: 14-seater minibus

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

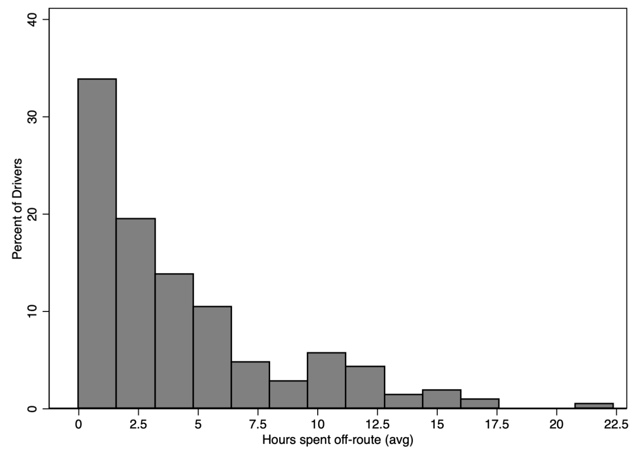
Figure A.5: Reckless Driving in the Control Group



(a) Speeding



(b) Braking

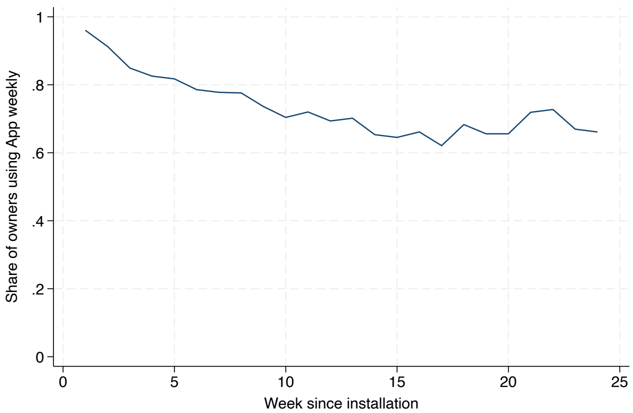


(c) Hours off-route

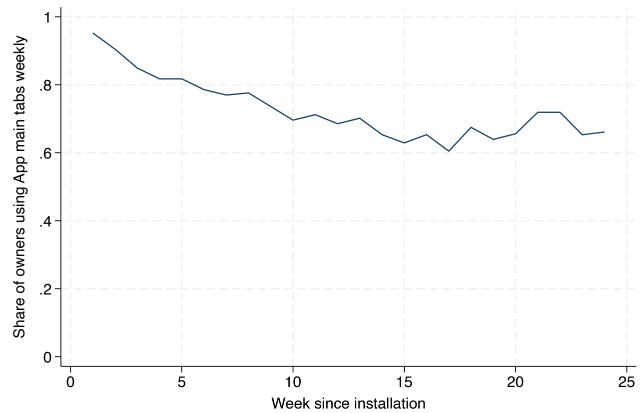
This figure presents the frequency of reckless driving in our sample of control matatus. Panel A: displays the share of days that drivers exceed 75km/h. Panel B: displays the share of days that drivers brake sharply. Panel C: displays the number of hours drivers spend off route ($>$ 400 meters from their designated route).

Figure A.6: App usage since installation

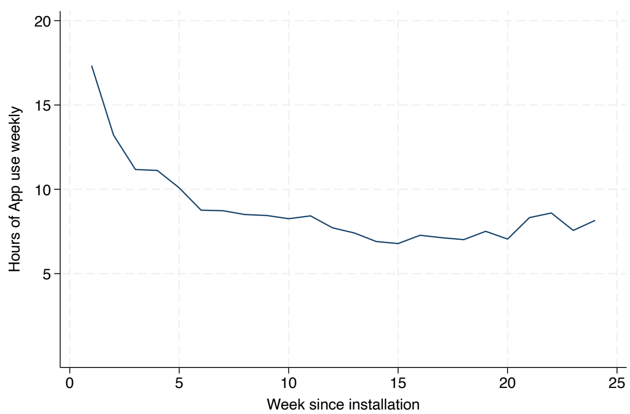
(a) Any Api Calls (weekly)



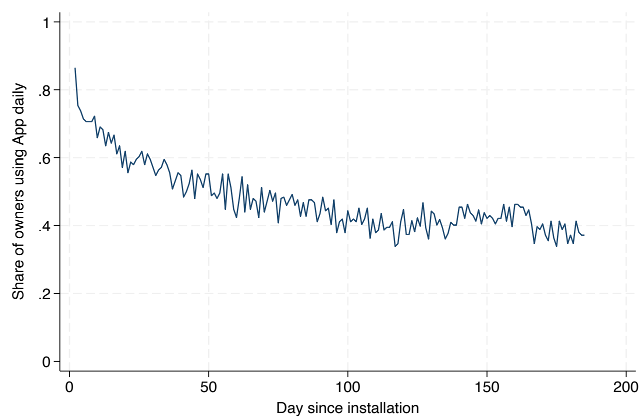
(b) Dashboard Api Calls (weekly)



(c) Hours of App use (weekly)

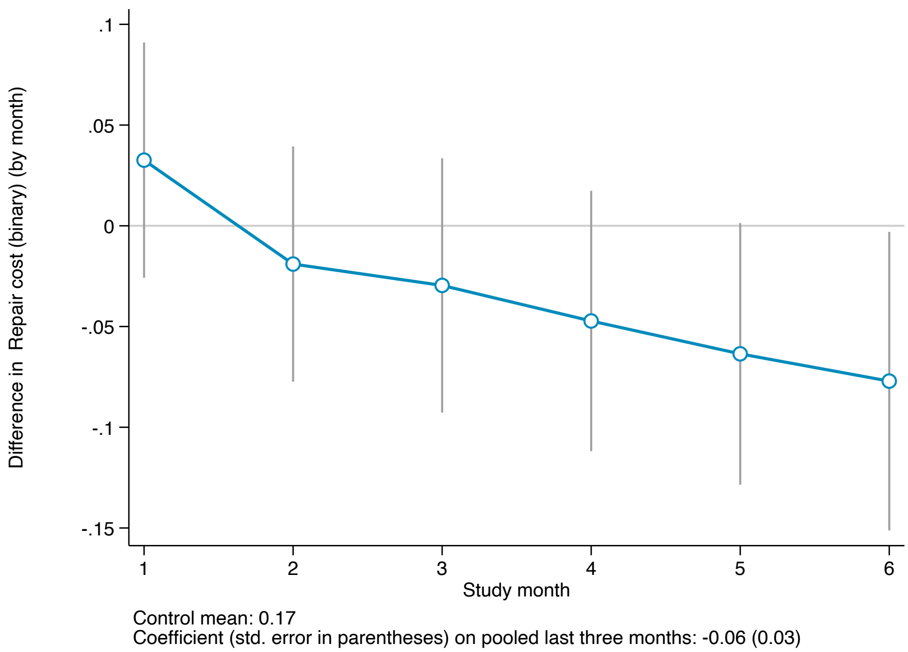


(d) Any Api Calls (daily)



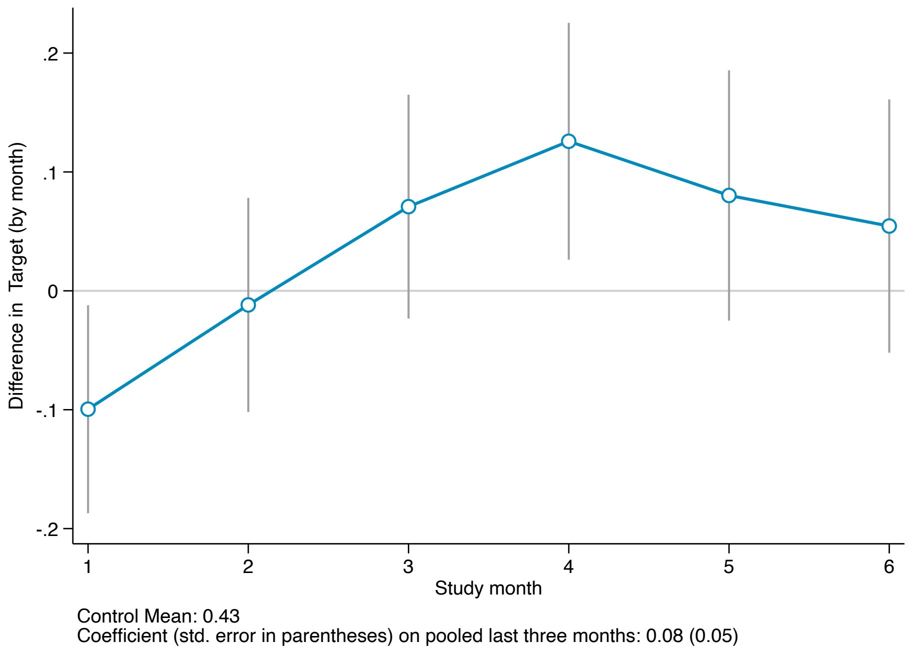
Notes: To measure device usage, we capture API calls that were made in a day. An API call is generated each time the owner requests data from the server, such as when logging in or refreshing a screen. Top left panel: the share of owners who make any API call by week. Top right panel: the share of owners who make any API to the main information dashboard tabs (excludes the survey) by week. Bottom left panel: the average number of hours owners use the App per week. Bottom right panel: the share of owners who make any API call by day.

Figure A.7: Treatment effects on large repair costs



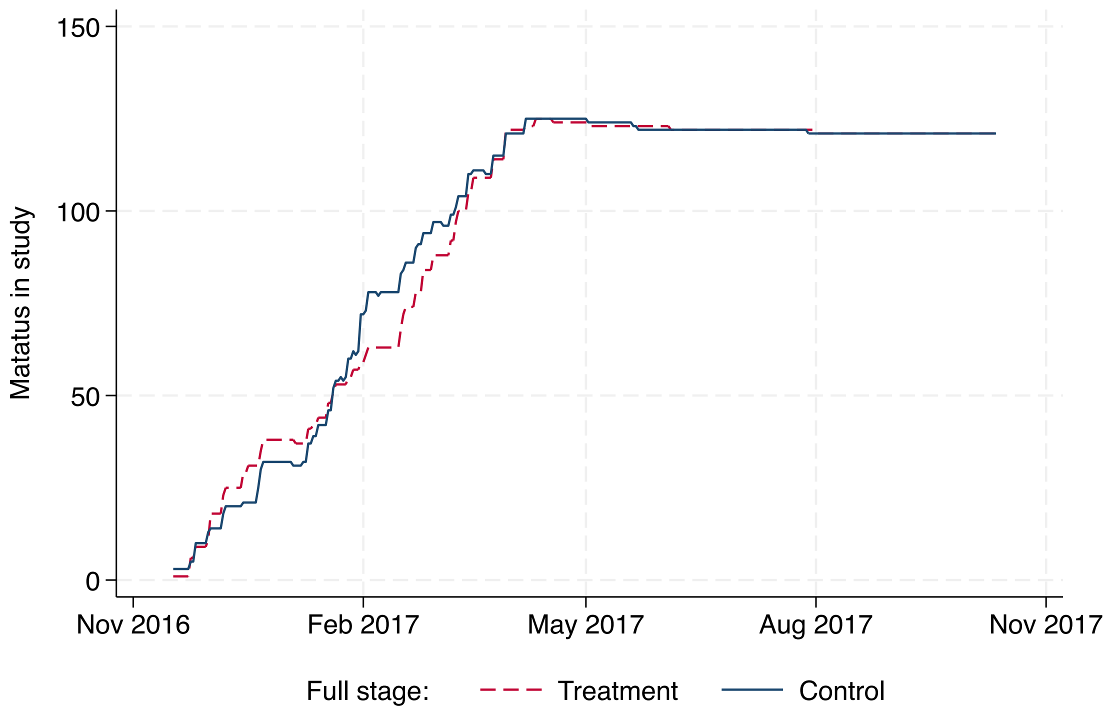
Notes: OLS estimates according to Equation 2. Treatment effect by month on probability of incurring a large repair cost (in excess of 1000 KES – the 85th percentile of repair costs). Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure A.8: Treatment effects on probability of making the target



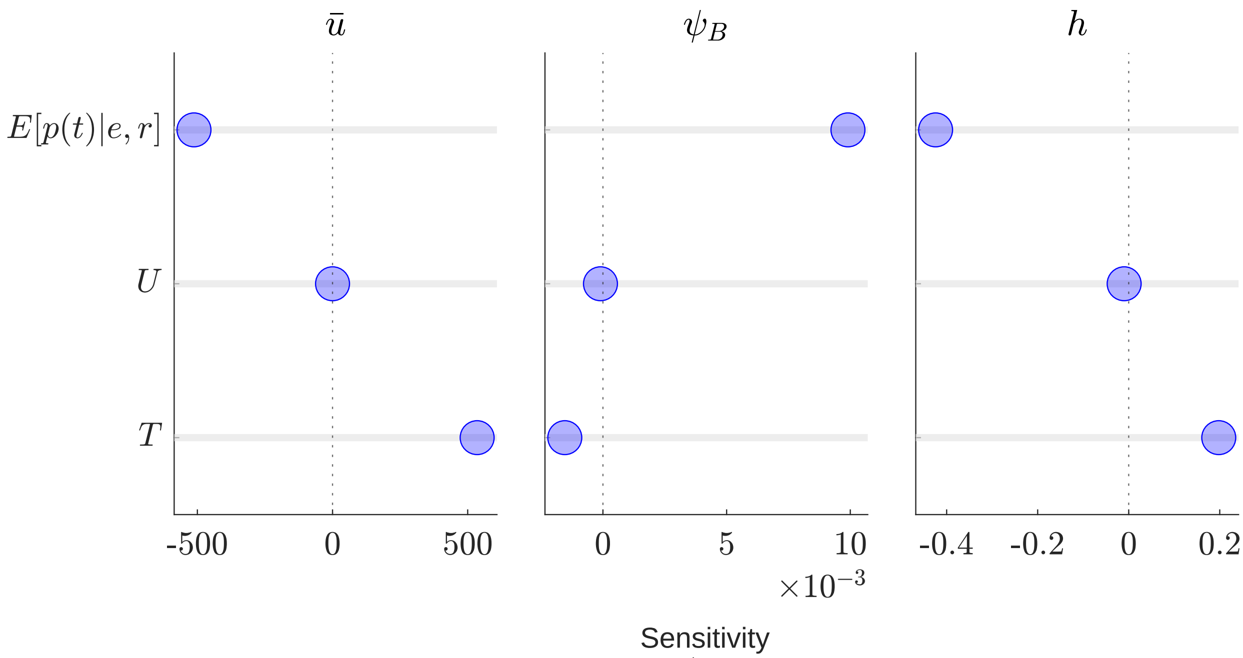
Notes: OLS estimates according to Equation 2. Treatment effect by month on probability of making the target. Standard errors for 95% confidence intervals clustered at the matatu level. We present the control group mean. We also present the coefficient (and standard error) of a regression of the outcome on an indicator for being in the last three months of the study (with same controls, fixed effects and standard errors as in Equation 2).

Figure A.9: Installation timeline



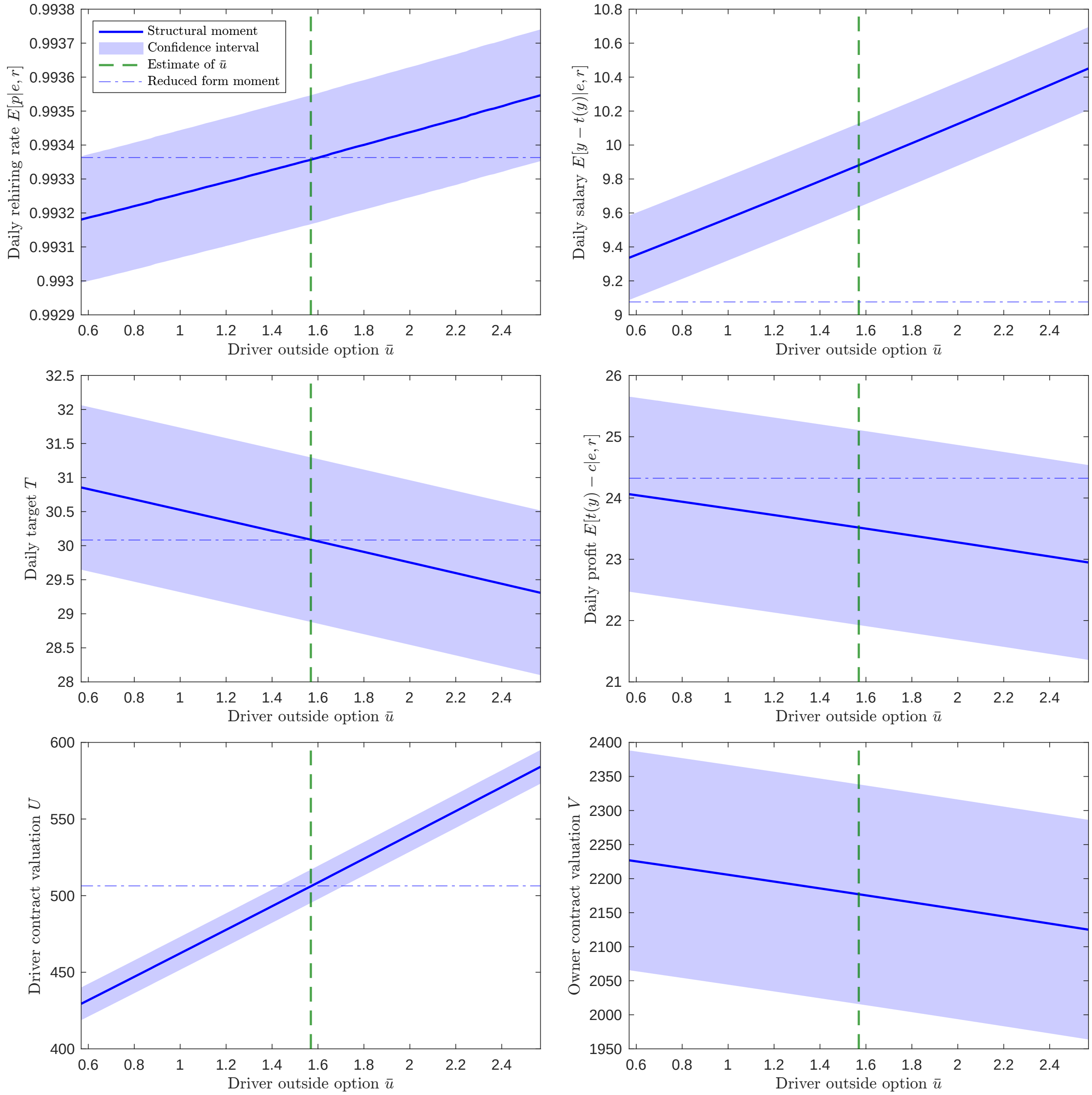
Notes: Number of matatus that were fitted with tracking devices (and hence were added to the study) per week. The first installation took place in November 2016, and continued until April 2017. On average, the field team was able to fit trackers to 15 matatus per week. As a result it took approximately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

Figure A.10: Estimate of sensitivity matrix Λ



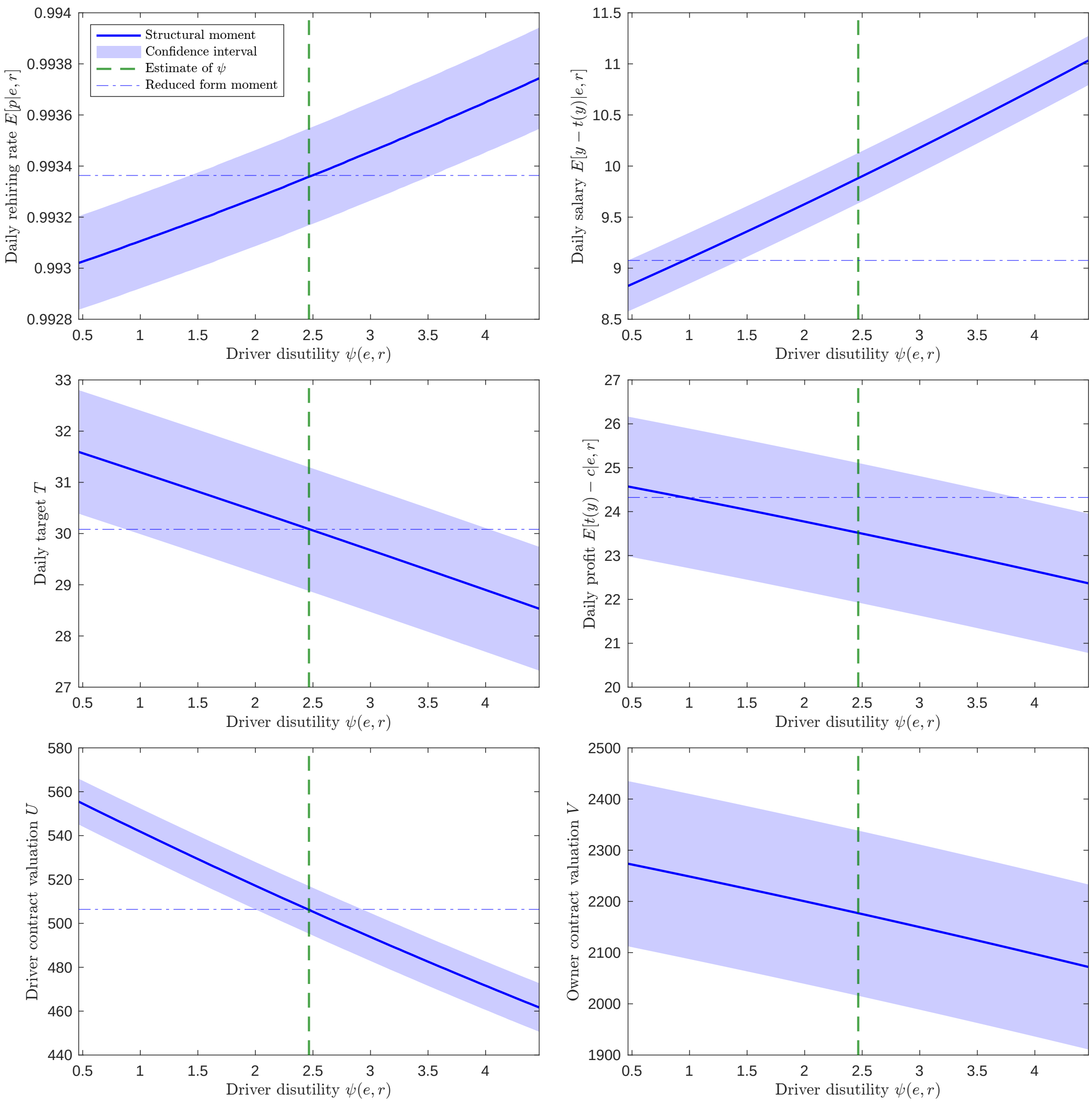
Notes: Figures plot estimated values of $\Lambda = (\mathbf{J}'\mathbf{W}\mathbf{J})^{-1}\mathbf{J}'\mathbf{W}$, where $\mathbf{J}$ is the 3×3 Jacobian matrix of derivatives of each of the three moments $\mathbb{E}[p(t)|e,r]$, U, and T with respect to each of the three parameters $\bar{u}$, ψ_B, and h; and $\mathbf{W}$ is a weighting matrix. Columns of Λ show the sensitivity in dollars of a given parameter estimate to a one-dollar change in each of the moments (i.e. the rows of Λ). See Appendix E for discussion.

Figure A.11: Model counterfactuals for changes in outside option $\bar{u}$



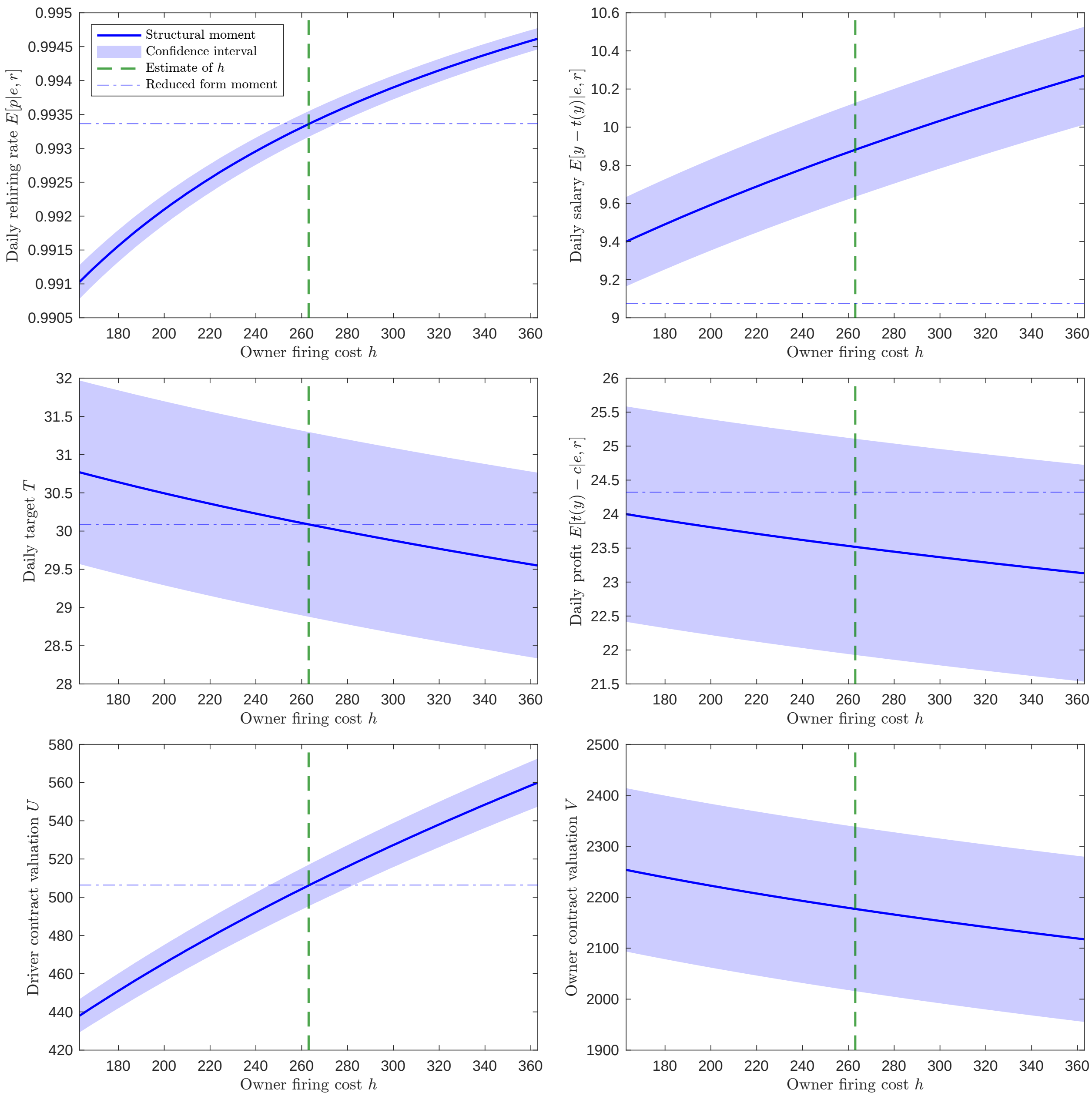
Notes: Figures plot counterfactual outcomes of interest for changes in the outside option $\bar{u}$. The three panels on the left (rehiring probability, target, and driver valuation) are the targeted moments, whereas the other three moments are untargeted. Dashed line shows the estimated outside option in the control group. The dash-dotted horizontal line shows reduced-form moments in the cross-section, where applicable (these may differ from the moments in the experimental estimates due to the lack weighting by number of bus-days). The other two parameters ($\psi(e, r)$ and h) are fixed at the values estimated in the control sample.

Figure A.12: Model counterfactuals for changes in driver disutility $\psi(e, r)$



Notes: Figures plot counterfactual outcomes of interest for changes in driver disutility $\psi(e, r)$. The three panels on the left (rehiring probability, target, and driver valuation) are the targeted moments, whereas the other three moments are untargeted. Dashed line shows estimated disutility in the control group. The dash-dotted horizontal line shows reduced-form moments in the cross-section, where applicable (these may differ from the moments in the experimental estimates due to the lack weighting by number of bus-days). The other two parameters ($\bar{u}$ and h) are fixed at the values estimated in the control sample.

Figure A.13: Model counterfactuals for changes in firing cost h



Notes: Figures plot counterfactual outcomes of interest for changes in firing cost h. The three panels on the left (rehiring probability, target, and driver valuation) are the targeted moments, whereas the other three moments are untargeted. Dashed line shows estimated disutility in the control group. The dash-dotted horizontal line shows reduced-form moments in the cross-section, where applicable (these may differ from the moments in the experimental estimates due to the lack weighting by number of bus-days). The other two parameters ($\bar{u}$ and $\psi(e,r)$) are fixed at the values estimated in the control sample.

C Appendix Tables

Table A.1: Differential reporting between treatment and control

	(1) Revenue	(2) Salary
Mileage in km	8.840	1.056
	(1.087)	(0.268)
Treatment	-259.196	-34.587
	(267.698)	(53.403)
Mileage X Treatment	1.242	0.383
	(1.487)	(0.397)
Control Mean of DV	7126.93	988.92
Day FE	Y	Y
Route FE	Y	Y
Matatu N	241	241
Matatu-Day N	22,436	22,426

Notes: This table shows how driver reports of revenue and salary differ by the number of miles driven and treatment. OLS regression of revenue (column 1) and salary (column 2) on the number of miles traveled, an indicator for treatment, and an interaction term between the two. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level.

Table A.2: Correlations with Target at Baseline

	(1) Target	(2) Target	(3) Target	(4) Target
Owner-Driver Tenure	1.52	1.74		
	(1.32)	(1.36)		
Value of Job		-0.00		
		(0.00)		
Driver Exp.		-3.11		
		(4.27)		
Driver Educ		-4.49		
		(37.61)		
Driver Risk		2.05		
		(9.01)		
Matatu Age			-17.45	
			(5.93)	
Number of Features			99.81	
			(28.95)	
Matatu Price				0.41
				(0.09)
Control Mean	3134	3134	3134	3134
Observations	256	256	255	256

Notes: This table presents correlations between the baseline target and various owner/driver/matatu characteristics. OLS regression of target on owner-driver tenure (column 1), driver characteristics (column 2), matatu characteristics (column 3), and matatu price (column 4). Driver characteristics in column 2 includes: value of job (represents drivers' willingness to accept to give up their job to another driver), driver experience (is the number of years the driver has worked in the industry), driver education (is their level of education), and driver risk (is the level of risk aversion we measure in a risk preferences game). Data from baseline surveys of owners and drivers. Robust standard errors.

Table A.3: Correlations with Business Outcomes in Control Group

	(1) Hours On	(2) Repair Costs	(3) Revenue	(4) Gross Profits
Own-Dri Tenure	-0.01	-5.17	-10.07	-1.76
	(0.01)	(3.81)	(4.87)	(7.23)
Value of Job	0.04	2.53	-3.83	-15.12
	(0.02)	(2.93)	(7.62)	(15.86)
Driver Exp.	0.20	3.46	15.30	2.44
	(0.08)	(17.09)	(23.88)	(32.88)
Driver Educ	-0.89	-170.47	-11.93	372.47
	(0.74)	(95.30)	(170.49)	(182.83)
Driver Risk Aversion	-0.10	-38.18	-4.93	45.75
	(0.13)	(16.14)	(45.89)	(44.62)
Control Mean	14	500	7090	3256
Observations	22891	7529	11205	5117

Notes: This table presents correlations between four core business outcomes and various owner/driver/matatu characteristics. OLS regression of hours the ignition was on (column 1), repair costs (column 2), revenue (column 3), and profits (column 4) on owner and driver characteristics, controlling for matatu characteristics. Owner-Driver tenure represents how long the owner and driver have worked together, value of job represents drivers' willingness to accept to give up their job to another driver, driver experience is the number of years the driver has worked in the industry, driver education is their level of education, and driver risk aversion is the level of risk aversion we measure in a risk preferences game. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data.

Table A.4: Treatment effects on effort, costs, revenue, target, profits, salary (pooled)

	(1) Device on (hours)	(2) Mileage (kilometers)	(3) Repair costs	(4) Repair costs (large)	(5) Revenue	(6) Target	(7) Met target	(8) Gross profit	(9) Salary per hour
Treatment	0.77	5.74	-107.2	-0.031	-71.6	-83.7	0.033	161.0	0.070
	(0.50)	(4.96)	(75.7)	(0.029)	(162.8)	(75.7)	(0.043)	(181.3)	(2.30)
Treatment × First 3 Months	0.40	2.98	-45.5	-0.0095	-133.6	-54.4	-0.0088	-9.57	0.47
	(0.50)	(5.09)	(70.3)	(0.029)	(176.0)	(79.0)	(0.043)	(193.8)	(2.42)
Treatment × Last 3 Months	1.11	8.32	-181.8	-0.057	-6.69	-118.9	0.082	366.7	-0.35
	(0.64)	(5.79)	(88.8)	(0.032)	(177.4)	(86.0)	(0.049)	(196.7)	(2.67)
Control Mean of DV	14.8	96.6	483.5	0.17	7126.9	3057.4	0.43	3260.5	61.3
Test Early = Late	0.22	0.25	0.01	0.02	0.36	0.32	0.01	0.02	0.71
Controls	X	X	X	X	X	X	X	X	X
Day FE	X	X	X	X	X	X	X	X	X
Route FE	X	X	X	X	X	X	X	X	X
Matatu-Day N	45,654	45,654	15,881	15,881	22,436	15,888	15,888	10,406	22,426

Notes: The first panel of this table presents treatment effects for all the experimental results (OLS regressions of outcome on an indicator for being treated and the standard controls and fixed effects as in Equation 2). The second panel of this table presents treatment effects for all the experimental results (OLS regressions of outcome on an indicator for being treated in the first three months, being treated last three months of the study, and the standard controls and fixed effects as in Equation 2). The omitted group is the control group. "Device on (hours)": number of hours the tracking device reported the ignition to be on. "Mileage (kilometers)": number of kilometers the tracking device reported the bus on the road. "Repair costs": owner-reported daily repair costs. "Large repair costs": owner-reported daily repair costs that exceed 10 USD. "Revenue": driver-reported daily revenue. "Target": daily revenue target set by owner. "Met target": whether the driver met the target. "Gross profit": Revenue minus repair costs minus driver residual claim (salary). "Salary": driver-reported residual claim (salary). 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, driver SMS reports, and aggregated tracking device data. We report the p-value from a test of whether the treatment effect in the first three months is equal to the treatment effect in the last three months.

Table A.5: Quantile Treatment Effects: Device On

	Quantile Treatment Effect: Device On				
	(1)	(2)	(3)	(4)	(5)
	10th	25th	50th	75th	90th
Treatment ×	-0.30	0.68	-0.73	-1.49	-2.01
Month 1	(1.00)	(1.20)	(0.33)	(0.38)	(0.31)
Treatment ×	0.23	1.91	0.33	-0.21	-0.71
Month 2	(1.06)	(0.99)	(0.33)	(0.32)	(0.29)
Treatment ×	1.91	1.95	0.32	-0.37	-0.57
Month 3	(1.65)	(0.93)	(0.34)	(0.33)	(0.31)
Treatment ×	0.52	1.28	0.20	0.03	-0.21
Month 4	(0.93)	(1.32)	(0.40)	(0.40)	(0.28)
Treatment ×	1.63	1.89	0.85	0.69	0.37
Month 5	(1.18)	(1.16)	(0.39)	(0.39)	(0.31)
Treatment ×	2.37	2.14	0.85	0.65	0.39
Month 6	(1.21)	(1.04)	(0.36)	(0.38)	(0.34)
Control Mean	14.63	14.63	14.63	14.63	14.63
Observations	45654	45654	45654	45654	45654

Notes: This table presents quantile treatment effects for the number of hours the device was on. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data.

Table A.6: Quantile Treatment Effects: Repair Costs

	Quantile Treatment Effect: Repair Cost				
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th
Treatment × Month 1	23.39 (22.28)	39.13 (25.42)	56.56 (28.57)	84.59 (106.13)	143.72 (193.28)
Treatment × Month 2	22.96 (24.75)	21.02 (26.79)	7.49 (27.98)	-84.86 (101.50)	-238.68 (259.80)
Treatment × Month 3	-10.32 (26.64)	-17.55 (26.71)	-50.08 (30.67)	-94.74 (96.87)	-305.06 (207.15)
Treatment × Month 4	-56.52 (27.60)	-42.92 (28.80)	-73.76 (27.97)	-151.03 (102.66)	-441.03 (178.41)
Treatment × Month 5	-19.93 (25.64)	-43.29 (24.65)	-85.91 (30.43)	-151.67 (111.11)	-453.90 (178.72)
Treatment × Month 6	-41.45 (24.41)	-55.23 (25.42)	-73.86 (29.38)	-185.02 (90.37)	-502.32 (154.35)
Control Mean	462.96	462.96	462.96	462.96	462.96
Observations	15886	15886	15886	15886	15886

Notes: This table presents quantile treatment effects for repair costs. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level.

Table A.7: Quantile Treatment Effects: Revenue

	Quantile Treatment Effect: Revenue				
	(1)	(2)	(3)	(4)	(5)
	10th	25th	50th	75th	90th
Treatment ×	-1345.63	-361.87	-137.66	-82.17	-198.48
Month 1	(801.43)	(232.84)	(159.20)	(152.87)	(151.38)
Treatment ×	-50.11	-111.02	3.41	-20.11	-207.08
Month 2	(509.28)	(198.18)	(155.98)	(137.08)	(137.11)
Treatment ×	48.10	-30.74	-25.29	-84.80	-323.00
Month 3	(419.53)	(237.85)	(145.20)	(147.84)	(142.55)
Treatment ×	367.48	14.49	14.08	-79.90	-201.69
Month 4	(386.56)	(263.02)	(152.34)	(152.12)	(193.61)
Treatment ×	154.05	26.87	31.44	-6.62	-269.74
Month 5	(401.01)	(264.64)	(152.62)	(150.40)	(145.64)
Treatment ×	81.80	-232.67	-119.18	-89.12	-304.78
Month 6	(308.74)	(235.16)	(160.12)	(149.34)	(154.63)
Control Mean	7096.65	7096.65	7096.65	7096.65	7096.65
Observations	22437	22437	22437	22437	22437

Notes: This table presents quantile treatment effects for revenue earned. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level.

Table A.8: Quantile Treatment Effects: Target

	Quantile Treatment Effect: Target				
	(1)	(2)	(3)	(4)	(5)
	10th	25th	50th	75th	90th
Treatment ×	-99.18	22.44	132.30	55.74	-112.83
Month 1	(95.04)	(82.38)	(61.18)	(74.48)	(126.70)
Treatment ×	-186.33	-21.42	64.86	58.13	-109.23
Month 2	(163.29)	(93.67)	(63.47)	(85.32)	(128.02)
Treatment ×	-165.45	-44.60	84.41	63.31	-116.71
Month 3	(121.81)	(124.54)	(66.52)	(77.61)	(133.03)
Treatment ×	-269.34	-62.85	85.51	61.85	-86.53
Month 4	(61.15)	(100.79)	(61.68)	(75.03)	(123.49)
Treatment ×	-260.26	-25.44	82.27	89.59	-70.39
Month 5	(80.22)	(104.62)	(67.88)	(79.74)	(125.84)
Treatment ×	-243.75	-32.03	115.93	106.63	-33.70
Month 6	(81.89)	(94.02)	(76.50)	(83.85)	(124.30)
Control Mean	3017.09	3017.09	3017.09	3017.09	3017.09
Observations	15893	15893	15893	15893	15893

Notes: This table presents quantile treatment effects for target. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level.

Table A.9: Quantile Treatment Effects: Profit

	Quantile Treatment Effect: Profit				
	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th
Treatment × Month 1	-358.19 (857.79)	-189.26 (251.25)	-122.52 (142.78)	-113.35 (116.90)	-122.09 (128.94)
Treatment × Month 2	504.95 (514.60)	-32.63 (288.94)	51.77 (135.48)	-2.71 (121.40)	26.48 (127.54)
Treatment × Month 3	669.26 (530.31)	78.66 (298.67)	36.92 (138.98)	-37.25 (125.55)	-30.24 (134.48)
Treatment × Month 4	945.91 (494.04)	353.60 (256.57)	191.49 (161.77)	154.26 (156.79)	150.11 (195.00)
Treatment × Month 5	766.75 (612.96)	458.63 (262.22)	239.83 (132.88)	134.33 (137.63)	134.66 (129.98)
Treatment × Month 6	720.60 (576.47)	238.19 (259.74)	-124.22 (151.14)	-100.85 (153.80)	-63.95 (154.43)
Control Mean	3307.15	3307.15	3307.15	3307.15	3307.15
Observations	10411	10411	10411	10411	10411

Notes: This table presents quantile treatment effects for profit. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Standard errors clustered at the owner/driver/matatu level.

Table A.10: Separation Rates

	(1) Separation
Treatment	-0.03 (0.05)
Control Mean	0.19
Observations	255

Notes: This table presents the rates separation between owners and drivers. OLS regression of whether the owner and driver separated during the study period on an indicator for treatment, standard controls, and route fixed effects. Data from daily panel collected from owner in app reports, driver SMS reports, and aggregated tracking device data. Data is collapsed to owner/driver/matatu level. Standard errors clustered at the owner/driver/matatu level.

Table A.11: Model estimation under baseline contract: higher subsistence

Panel A: Assumptions

Input	Value	Notes
Subsistence income w	8	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on control group
Discount factor δ	0.99	

Panel B: GMM parameter estimates

Input	Value	Interpretation
Baseline driver disutility $\psi(e_B, r_B)$	1.47	Driving disutility of \$1.47
	(0.67)	
Firing cost h	279	Lost profit of firing of \$279
	(47)	(about 12 days of profit)
Driver outside option $\bar{u}$	1.12	Similar to unskilled daily wage with
	(0.38)	subsistence (\$1.12 + \$8 = \$9.12)

Panel C: Reduced form, structural, and matched moments

Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Firing probability $E[p\|e, r]$	0.007	0.007	0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	505.8	0.5
	(12.9)	(67.2)	(68.5)
Target T	30.1	30.1	-0.0
	(0.4)	(0.7)	(0.8)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	9.1	9.2	-0.1
	(0.2)	(0.5)	(0.5)
Owner profit $E[t\|e, r] - E[c\|e, r]$	23.3	23.2	0.1
	(0.7)	(0.5)	(0.9)
Owner contract value V	—	2,137	—
	—	(18)	—
Welfare $U + V$	—	2,643	—
	—	(70)	—

Notes: Generalized method of moments (GMM) estimation of driver disutility, firing cost, and outside option. Unlike in the main tables, we assume a higher subsistence income of \$8 (instead of \$7). Sample: control group. Targeted moments: Separation probability, driver contract value, and target. "Reduced form" as observed in the sample. "Structural" are the corresponding estimated model predictions. The difference is between reduced form and structural moments. Standard errors of parameters based on estimated asymptotic variance and of structural moments via the Delta Method.

Table A.12: Reduced form versus structural treatment estimation: higher subsistence

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	8	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on treated group
Discount factor δ	0.99	
Outside option $\bar{u}$	1.12	Estimated in control group
Firing cost h	279	Estimated in control group

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Disutility with monitoring $\psi(e_M, r_M)$	2.76	Increase of $1.29 (88%)
	(2.24)	

Panel C: Reduced form, structural, and matched treatment effects			
Treatment effect	Red. form (Δ)	Structural (Δ)	Difference (Δ)
Targeted moments:			
Firing probability $E[p\|e, r]$	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Target T	-1.0	-1.1	0.2
	(0.6)	(1.0)	(1.2)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	0.1	0.9	-0.8
	(0.2)	(0.5)	(0.6)
Owner profit $E[t\|e, r] - E[c\|e, r]$	1.7	0.8	0.8
	(1.0)	(1.2)	(1.6)
Driver contract value U	—	-21.7	—
	—	(22.0)	—
Owner contract value V	—	85.9	—
	—	(120.9)	—
Welfare $U + V$	—	64.3	—
	—	(125.1)	—

Notes: Driver disutility under monitoring estimated via GMM. Unlike in the main tables, we assume a higher subsistence income of $8 (instead of $7). Sample: treatment group. Targeted moments: Separation probability and target. Untargeted moments: driver contract value, driver salary, owner profit. "Reduced form (Δ)" are the difference between the treatment group and the control group in the data. "Structural (Δ)" are the corresponding difference between estimated model predictions of the treatment and control groups. "Difference (Δ)" is the difference between between reduced form and structural moment differences. Standard errors via the bootstrap.

Table A.13: Model estimation under baseline contract: lower subsistence

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	6	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on control group
Discount factor δ	0.99	

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Baseline driver disutility $\psi(e_B, r_B)$	3.47	Driving disutility of \$3.47
	(0.64)	
Firing cost h	247	Lost profit of firing of \$247
	(94)	(about 10 days of profit)
Driver outside option $\bar{u}$	1.98	Similar to unskilled daily wage with
	(0.32)	subsistence (\$1.98 + \$6 = \$7.98)

Panel C: Reduced form, structural, and matched moments			
Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Firing probability $E[p\|e, r]$	0.007	0.007	-0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	508.3	-2.0
	(12.9)	(64.0)	(65.2)
Target T	30.1	30.1	0.0
	(0.4)	(0.6)	(0.8)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	9.1	10.6	-1.6
	(0.2)	(0.5)	(0.5)
Owner profit $E[t\|e, r] - E[c\|e, r]$	25.3	23.8	1.6
	(0.7)	(0.5)	(0.9)
Owner contract value V	—	2,214	—
	—	(48)	—
Welfare $U + V$	—	2,722	—
	—	(80)	—

Notes: Generalized method of moments (GMM) estimation of driver disutility, firing cost, and outside option. Unlike in the main tables, we assume a lower subsistence income of \$6 (instead of \$7). Sample: control group. Targeted moments: Separation probability, driver contract value, and target. "Reduced form" as observed in the sample. "Structural" are the corresponding estimated model predictions. The difference is between reduced form and structural moments. Standard errors of parameters based on estimated asymptotic variance and of structural moments via the Delta Method.

Table A.14: Reduced form versus structural treatment estimation: lower subsistence

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	6	Kink in transfer schedule
Revenue distribution $G(\cdot\|e,r)$	—	Normal distribution on treated group
Discount factor δ	0.99	
Outside option $\bar{u}$	1.98	Estimated in control group
Firing cost h	247	Estimated in control group

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Disutility with monitoring $\psi(e_M, r_M)$	4.69	Increase of $1.22 (35%)
	(1.82)	

Panel C: Reduced form, structural, and matched treatment effects			
Treatment effect	Red. form (Δ)	Structural (Δ)	Difference (Δ)
Targeted moments:			
Firing probability $E[p\|e,r]$	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)
Target T	-1.0	-1.1	0.2
	(0.6)	(1.1)	(1.3)
Untargeted moments:			
Driver salary $E[y\|e,r] - E[t\|e,r]$	0.1	0.9	-0.9
	(0.2)	(0.8)	(0.8)
Owner profit $E[t\|e,r] - E[c\|e,r]$	1.7	0.8	0.9
	(1.0)	(1.2)	(1.6)
Driver net contract value U	—	-16.3	—
	—	(19.1)	—
Owner contract value V	—	82.4	—
	—	(121.7)	—
Welfare $U + V$	—	66.2	—
	—	(122.2)	—

Notes: Driver disutility under monitoring estimated via GMM. Unlike in the main tables, we assume a lower subsistence income of $6 (instead of $7). Sample: treatment group. Targeted moments: Separation probability and target. Untargeted moments: driver contract value, driver salary, owner profit. "Reduced form (Δ)" are the difference between the treatment group and the control group in the data. "Structural (Δ)" are the corresponding difference between estimated model predictions of the treatment and control groups. "Difference (Δ)" is the difference between between reduced form and structural moment differences. Standard errors via the bootstrap.

Table A.15: Model estimation under baseline contract: higher discount rate

Panel A: Assumptions

Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on control group
Discount factor δ	0.992	

Panel B: GMM parameter estimates

Input	Value	Interpretation
Baseline driver disutility $\psi(e_B, r_B)$	3.48	Driving disutility of \$3.48
	(0.66)	
Firing cost h	224	Lost profit of firing of \$224
	(15)	(about 9 days of profit)
Driver outside option $\bar{u}$	1.27	Similar to unskilled daily wage with
	(0.29)	subsistence (\$1.27 + \$7 = \$8.27)

Panel C: Reduced form, structural, and matched moments

Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Firing probability $E[p\|e, r]$	0.007	0.007	-0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	507.3	-0.9
	(12.9)	(82.3)	(83.3)
Target T	30.1	30.1	0.0
	(0.4)	(0.7)	(0.8)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	9.1	9.9	-0.8
	(0.2)	(0.5)	(0.5)
Owner profit $E[t\|e, r] - E[c\|e, r]$	24.3	23.5	0.8
	(0.7)	(0.5)	(0.9)
Owner contract value V	—	2,753	—
	—	(14)	—
Welfare $U + V$	—	3,260	—
	—	(83)	—

Notes: Generalized method of moments (GMM) estimation of driver disutility, firing cost, and outside option. Unlike in the main tables, we assume a higher discount rate of 0.992 (instead of 0.99). Sample: control group. Targeted moments: Separation probability, driver contract value, and target. "Reduced form" as observed in the sample. "Structural" are the corresponding estimated model predictions. The difference is between reduced form and structural moments. Standard errors of parameters based on estimated asymptotic variance and of structural moments via the Delta Method.

Table A.16: Reduced form versus structural treatment estimation: higher discount rate

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e,r)$	—	Normal distribution on treated group
Discount factor δ	0.992	
Outside option $\bar{u}$	1.27	Estimated in control group
Firing cost h	224	Estimated in control group

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Disutility with monitoring $\psi(e_M, r_M)$	4.72	Increase of $1.24 (35%)
	(2.09)	

Panel C: Reduced form, structural, and matched treatment effects			
Treatment effect	Red. form (Δ)	Structural (Δ)	Difference (Δ)
Targeted moments:			
Firing probability $E[p\|e,r]$	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)
Target T	-1.0	-1.1	0.2
	(0.6)	(1.2)	(1.4)
Untargeted moments:			
Driver salary $E[y\|e,r] - E[t\|e,r]$	0.1	0.9	-0.9
	(0.2)	(0.9)	(0.9)
Owner profit $E[t\|e,r] - E[c\|e,r]$	1.7	0.8	0.9
	(1.0)	(1.3)	(1.7)
Driver contract value U	—	-19.9	—
	—	(22.3)	—
Owner contract value V	—	103.5	—
	—	(161.3)	—
Welfare $U + V$	—	83.6	—
	—	(165.0)	—

Notes: Driver disutility under monitoring estimated via GMM. Unlike in the main tables, we assume a higher discount rate of 0.992 (instead of 0.99). Sample: treatment group. Targeted moments: Separation probability and target. Untargeted moments: driver contract value, driver salary, owner profit. "Reduced form (Δ)" are the difference between the treatment group and the control group in the data. "Structural (Δ)" are the corresponding difference between estimated model predictions of the treatment and control groups. "Difference (Δ)" is the difference between between reduced form and structural moment differences. Standard errors via the bootstrap.

Table A.17: Model estimation under baseline contract: lower discount rate

Panel A: Assumptions

Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e,r)$	—	Normal distribution on control group
Discount factor δ	0.988	

Panel B: GMM parameter estimates

Input	Value	Interpretation
Baseline driver disutility $\psi(e_B, r_B)$	1.46	Driving disutility of \$1.46
	(0.65)	
Firing cost h	299	Lost profit of firing of \$299
	(19)	(about 12 days of profit)
Driver outside option $\bar{u}$	1.87	Similar to unskilled daily wage with
	(0.43)	subsistence (\$1.87 + \$7 = \$8.87)

Panel C: Reduced form, structural, and matched moments

Control group outcome	Reduced form	Structural	Difference
Targeted moments:			
Firing probability $E[p\|e,r]$	0.007	0.007	-0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	507.9	-1.5
	(12.9)	(54.3)	(55.8)
Target T	30.1	30.1	0.0
	(0.4)	(0.7)	(0.8)
Untargeted moments:			
Driver salary $E[y\|e,r] - E[t\|e,r]$	9.1	9.9	-0.8
	(0.2)	(0.5)	(0.5)
Owner profit $E[t\|e,r] - E[c\|e,r]$	24.3	23.5	0.8
	(0.7)	(0.5)	(0.9)
Owner contract value V	—	1,794	—
	—	(11)	—
Welfare $U + V$	—	2,302	—
	—	(55)	—

Notes: Generalized method of moments (GMM) estimation of driver disutility, firing cost, and outside option. Unlike in the main tables, we assume a lower discount rate of 0.988 (instead of 0.99). Sample: control group. Targeted moments: Separation probability, driver contract value, and target. "Reduced form" as observed in the sample. "Structural" are the corresponding estimated model predictions. The difference is between reduced form and structural moments. Standard errors of parameters based on estimated asymptotic variance and of structural moments via the Delta Method.

Table A.18: Reduced form versus structural treatment estimation: lower discount rate

Panel A: Assumptions		
Input	Value	Notes
Subsistence income w	7	Kink in transfer schedule
Revenue distribution $G(\cdot\|e, r)$	—	Normal distribution on treated group
Discount factor δ	0.988	
Outside option $\bar{u}$	1.87	Estimated in control group
Firing cost h	299	Estimated in control group

Panel B: GMM parameter estimates		
Input	Value	Interpretation
Disutility with monitoring $\psi(e_M, r_M)$	2.79	Increase of \$1.33 (91%)
	(3.00)	

Panel C: Reduced form, structural, and matched treatment effects			
Treatment effect	Red. form (Δ)	Structural (Δ)	Difference (Δ)
Targeted moments:			
Firing probability $E[p\|e, r]$	0.000	-0.000	0.000
	(0.000)	(0.001)	(0.001)
Target T	-1.0	-1.1	0.2
	(0.6)	(1.2)	(1.3)
Untargeted moments:			
Driver salary $E[y\|e, r] - E[t\|e, r]$	0.1	0.9	-0.9
	(0.2)	(0.7)	(0.8)
Owner profit $E[t\|e, r] - E[c\|e, r]$	1.7	0.8	0.9
	(1.0)	(1.3)	(1.7)
Driver contract value U	—	-21.4	—
	—	(17.8)	—
Owner contract value V	—	68.6	—
	—	(110.9)	—
Welfare $U + V$	—	47.2	—
	—	(114.1)	—

Notes: Driver disutility under monitoring estimated via GMM. Unlike in the main tables, we assume a lower discount rate of 0.988 (instead of 0.99). Sample: treatment group. Targeted moments: Separation probability and target. Untargeted moments: driver contract value, driver salary, owner profit. "Reduced form (Δ)" are the difference between the treatment group and the control group in the data. "Structural (Δ)" are the corresponding difference between estimated model predictions of the treatment and controls groups. "Difference (Δ)" is the difference between between reduced form and structural moment differences. Standard errors via the bootstrap.

D Model Details

D.1 Owner Problem

The fully specified owner problem is given by

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

subject to

1. $U - \bar{u} = \mathbb{E}\left[y - t(y) + p(t(y))\left(\delta U - \bar{u}\right)|e, r\right] - \psi\left(e, r\right) \geq 0$

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

3. $t(y) \leq y$

4. $y - t(y) + p(t(y))\left(\delta U - \bar{u}\right) \geq y$

5. $t(y) \in \arg\max_{\tilde{t}\geq 0} y - \tilde{t} + p(\tilde{t})\left(\delta U - \bar{u}\right)$,

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. 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.[51]

[51] 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.

D.2 Assumptions

D.2.1 Technology and Preferences

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* $g(y|e,r)$ *and is from the location-scale family (to ensure the distribtution is well-defined after subtracting subsistence).*

- *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* $s \in \{e,r\}$.

- *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* $s \in \{e,r\}$.

D.2.2 Relative Preferences for Risk

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*

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

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:

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

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.

D.3 Proofs

D.3.1 Lemmas

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

$$p(t) = 1 - \frac{T - t}{\delta U - \bar{u}}$$

for $t \in (T - (\delta U - \bar{u}), T)$ *for some* $T > 0$.

Proof. The proof proceeds in two steps: in step 1, we show that under Assumption 1 the minimally optimal slope of $p(t)$ is $\frac{1}{\delta U - \bar{u}}$ – that is, a $p(t)$ with lower slope than $\frac{1}{\delta U - \bar{u}}$ 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:

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

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, $p(y)\left(\delta U - \bar{u}\right) \geq y$ 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 $p(t)\left(\delta U - \bar{u}\right)$. We now split the owner's objective function $\mathbb{E}\left[y - t + p(t)\left(\delta U - \bar{u}\right)|e, r\right]$ into intervals $\mathcal{T}$ and $\mathcal{X}_i$ for $i = 1, ..., I$:

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

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

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

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 $\mathcal{T}$, $p(y)(\delta U - \bar{u}) \geq y$. Together with the MLRP, this implies that

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

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 $p(t) \geq 1 - \frac{T-t}{\delta U - \bar{u}}$) as:

$$\begin{aligned} X(e,r) &= \mathbb{E}\left[t - c(r) + \delta V - (1-p(t))h | e, r\right] &(3)\\ &= \int_0^T \left[y - (1-p(y))h\right] g\left(y|e,r\right) dy + \left(1 - G\left(T|e,r\right)\right) T + \delta V - \mathbb{E}\left[c(r)|r\right] &(4)\end{aligned}$$

and the corresponding marginal effect of effort and risk:

$$X_s\left(e,r\right) = \int_0^T \left[y - (1-p(t))h\right] g_s\left(y|e,r\right) dy - G_s\left(T|e,r\right) T - \frac{\partial \mathbb{E}\left[c(r)|r\right]}{\partial s}.$$

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 $g_s\left(y|e,r\right) \leq 0$ for all $y \leq y^*$, and (c) $h = 0$. 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

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

with $\phi_s = -\frac{\delta}{1-\delta} G_s \left(y^* | e, r\right)$. 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 $(e,r) \geq (e_D, r_D)$:

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

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

□

D.3.2 Propositions

The social planner ("owner-driver") problem is

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

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: $\mathbb{E}\left[y - c(r_B) | e_B, r_B\right] - \psi\left(e_B, r_B\right) < \mathbb{E}\left[y - c(r^*) | e^*, r^*\right] - \psi\left(e^*, r^*\right)$.*

Proof of Proposition 1

Proof. Using the Lemma, we can write the driver's problem as: $\delta U - T + \mathbb{E}\left[y | e, r\right] - \psi(e, r)$. 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:

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

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.

$\square$

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:*

$$p(t,e,r) = \begin{cases} 1 - \frac{T-t}{\delta U_M - \bar{u}} & \text{if } e = e_M \text{ and } r = r_M \\ 0 & \text{otherwise} \end{cases}$$

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: $\mathbb{E}[y|e_M, r_M] \lesseqgtr \mathbb{E}[y|e_B, r_B]$.*
   - *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 $\mathbb{E}[y|e_M, r_M] \leq \mathbb{E}[y|e_B, r_B]$.*
   - *The welfare effect is ambiguous:*

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

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

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

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 $p(0,e,r) + \frac{y}{\delta U - \bar{u}} < 1$ and hence $\hat{T} = \inf\{t : \hat{p}(t,e,r) = 1\}$. For every such rehiring schedule, there exists another with equal expected rehiring probability but at minimal slope, i.e. $\mathbb{E}\left[1 - \frac{T-t}{\delta U - \bar{u}}|e,r\right] = \mathbb{E}\left[\hat{p}(t,e,r)|e,r\right]$ with $T > \hat{T}$. 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 $U_M \geq \mathbb{E}[y|e, r] - \psi(e, r)$. 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

$$p(t, e, r) = \begin{cases} 1 - \frac{T-t}{\delta U_M} & \text{if } e = e_M \text{ and } r = r_M \\ 0 & \text{otherwise} \end{cases}$$

is a solution to the problem.

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

$$\max_{(e,r)\in\mathcal{S}, T\in\mathcal{Y}} \delta V + T - G\left(T|e, r\right)\left\{T - \mathbb{E}\left[y|e, r, y \leq T\right]\right\}\left(1 + \frac{h}{\delta U - \bar{u}}\right) - \mathbb{E}\left[c(r)|r\right]. \qquad (5)$$

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:

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

where $U(e, r, T) = \left(\mathbb{E}[y|e, r] - \psi(e, r) - T\right)/(1 - \delta)$. 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 $\mathbb{E}[y|e, r]$ 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 (5) with respect to T is:

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

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

However, in case revenue falls, $\mathbb{E}[y|e_M, r_M] \leq \mathbb{E}[y|e_B, r_B]$, 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.

□

D.4 Model with Risk Signal

One simplification in the main model is that we assume that owners do not receive a signal about the level of risk. However, it is plausible that owners may receive a signal about the risk r taken by the driver from the repair costs incurred, $c \sim F(\cdot|r)$. We explore the implications of this more general version of the model in this section.

The basic intuition of our model extension is as follows. When the owner pursues a grim trigger strategy, i.e. firing the driver when repair costs are above some threshold, the degree to which drivers are sensitive to this threat depends critically on how strongly their actions impact repair costs relative to other factors (i.e. the signal-to-noise ratio). We first use our data to estimate the magnitude of the signal-to-noise ratio of risk choice to repair costs, which we find to be very low (repair costs are affected by much more than risky driving). Second, we can estimate how much drivers adjust their risk behavior in light of the grim trigger threat and this signal-to-noise ratio. We find that driver risk behavior only shifts by 2.7% towards the owner's preferred risk choice. In contrast, under monitoring, the owner is able to shift the driver entirely to their preferred effort-risk profile because they directly observe the behavior they want to control rather than a poor proxy of it. In other words, as long as repair costs are a noisy signal of risk taking there would remain benefits from monitoring - both by allowing precise contracting on r and by eliminating the inefficient firings necessary to maintain these incentives (i.e. firing an unlucky driver when repair costs are high even if the driver chose the "correct" r). As a result, risk observability is not sufficient to meaningfully shift the driver's risk choice towards the owner's preferred choice, whereas monitoring can lead to a substantial shift.

More formally, to generalize the model along these lines, we extend the baseline rehiring function $p(t)$ to be a function of both the transfer t and the repair cost c, $p(t, c)$. In doing so, the owner may be able to structure the contract with the driver in a way that incentivizes the driver to reduce risk, similarly to the contract under monitoring. We now explore the assumptions that must hold for repair cost observability to generate a meaningful reduction in risk. These assumptions turn out to be implausibly strong in our context. We then quantify how close to the first-best outcome an owner may get with a noisy signal of risk, showing that it would likely close only a small part of the gap between the baseline contract and the owner's preferred contract under monitoring.

First, notice that the Lemma still applies with respect to the transfer: that is, to provide incentives to the driver to transfer revenue to the owner, the owner has to satisfy the minimal slope of $1/(\delta U - \bar{u})$ with respect to t; and steeper slopes would incentivize both higher effort *and* higher risk. Hence, we can write the rehiring function as being additively separable between the transfer and the cost:

$$p(t, c) = p_0 + p_1(t) + p_2(c),$$

where $p_1(t) = t/(\delta U - \bar{u})$ for all values of t and c such that $0 \leq p(t, c) \leq 1$. We now define a cost cutoff C that, if the transfer t matches the target T, guarantees rehiring: $p(T, C) = 1$. Solving for p_0, we then get

$$p(t, c) = 1 - \frac{T - t}{\delta U - \bar{u}} - b(c)$$

where $b(c) = p_2(C) - p(c)$, which we call the *cost-specific firing function*. To see how this function affects driving incentives, we can enter $p(t, c)$ into the driver's valuation, where the second line follows after simplification:

$$
\begin{aligned}
U &= \mathbb{E}\left[y - t(y) + p(t,c)\left(\delta U - \bar{u}\right)|e,r\right] - \psi(e,r) + \bar{u} \\
&= \mathbb{E}\left[y|e,r\right] + \left(1 - \mathbb{E}\left[b(c)|r\right]\right)\left(\delta U - \bar{u}\right) - T - \psi(e,r) + \bar{u},
\end{aligned}
$$

which shows that the driver's continuation value falls by the expected cost-specific firing rate. For any given risk choice r, the driver considers the probability he will be fired, given a firing function $b(c)$, which may affect his choice of r. This affects the incentive-compatible set of effort-risk choices as follows:

$$
\mathcal{I}(b(c)) = \left\{(e,r) \in \mathcal{S} : \frac{\int_0^{\bar{y}} y g_e(y|e,r)dy}{\int_0^{\bar{y}} y g_r(y|e,r)dy - \left(\delta U - \bar{u}\right)\int_0^{\bar{c}} b(c) f_r(c|r)dc} = \frac{\psi_e(e,r)}{\psi_r(e,r)}\right\}
$$

where $f_r(c|r) = \partial f(c|r)/\partial r$ and $f(c|r)$ is the density associated with $F(c|r)$. This modified incentive-compatible set shows that the driver's optimal effort-risk choice will be more strongly geared towards effort and away from risk. To illustrate, in Figure 2, panel (a), the line representing the incentive-compatible set would be flatter. This confirms that making rehiring conditional on observed cost does indeed allow the owner to incentivize effort-risk choices that are closer to her first-best (as well as the choices under monitoring). We now turn to the question of how close she may get.

To examine the strongest possible risk-reducing incentive the owner could possibly provide (whether or not this would be optimal), we examine the grim trigger strategy $b(c) = 1\left[c \geq \hat{c}\right]$, where $\hat{c} \in (0, \bar{c})$ is the trigger point. In this case, the expected firing function becomes $\mathbb{E}\left[b(c)|r\right] = 1 - F(\hat{c}|r)$. The marginal incentive with respect to risk is then:

$$
\frac{\partial U}{\partial r} = \int_0^{\bar{y}} y g_r(y|e,r)dy + F_r(\hat{c}|r)\left(\delta U - \bar{u}\right).
$$

Note that $F_r(\hat{c}|r)$ is smaller as the noise in costs c increases. To make this concrete, assume that $F(c|r)$ is Normal with mean $\mu(r)$, which is a continuously differentiable function of risk with derivative $\mu'(r)$, and variance σ^2. The marginal incentive is then

$$
F_r(\hat{c}|r) = -\frac{\mu'(r)}{\sigma}\phi\left(\frac{\hat{c} - \mu(r)}{\sigma}\right),
$$

where $\phi(\cdot)$ is the pdf of a Standard Normal random variable. As costs become less informative as a signal for risk, i.e. $\sigma \to \infty$, marginal incentives go to zero. To roughly quantify the magnitude of this possible maximum incentive in our setting, we assume the trigger point is at the mean of the distribution (which maximizes the marginal incentive), $\phi(0) \approx 0.4$.[52] We estimate σ to be about KES 1,067 across our daily panel. To estimate $\mu'(r)$, we regress repair costs and our measure of risk aversion measured at baseline and expressed in standard deviations, which yields a coefficient of KES 67 (SE KES 8) increase in repair costs for a standard-deviation increase in driver risk prefernce. Putting these estimates together, we arrive at $\widehat{F}_r(\hat{c}|r) = -0.027$. While this estimate relies on a number of assumptions—Normality of costs and risk aversion being a good proxy for estimating the risk-cost relationship—they illustrate that a grim trigger strategy on its own is unlikely to be quantitatively meaningful in our setting.

In addition to the strength of incentives for drivers, by making firing a function of costs, the owner also risks having to pay the hiring cost h even when a driver chose a low level of risk but had a high draw of costs from the cost distribution. This additional cost dissuades owners from using the strongest cost incentives in their rehiring decisions. In sum, given the weak ability to dissuade risk taking and the direct cost to owners from more firing, even if owners do contract on c this is unlikely to meaningfully change the value of the contract to either party or to change the implications of monitoring.

These conclusions are also consistent with our experience from conversations with owners about cost as a signal of driver risk taking. They point out that it takes months for break pads or shock absorbers to show a reliable sign of bad driving behavior. Only in the most egregious cases—such as major accidents—do owners who do not have access to monitoring typically fire drivers based on repair cost.

[52] Note that this would imply that the owner would fire the driver 50% of the time since the trigger is set at the mean, which is also the median, of the cost distribution for the equilibrium choice of risk. So the owner would prefer to set the trigger at a much higher risk, which would then be associated with an even smaller value of $\phi(\cdot)$. But setting the trigger point at the mean provides an informative upper bound for the strength of driver incentives.

E Structural Estimation Details

The goals of the structural estimation are twofold: first, we seek to quantify the valuation of the relational contract between the owner and the driver; and second, we aim to estimate the welfare effects of introducing monitoring into this relationship. We begin by discussing identification and calibration of the model, after which we discuss estimation.

E.1 Identification and Calibration

There are four parameters in the model: $\{\bar{u}, \psi_B, \psi_M, h\}$. $\bar{u}$ is the daily outside option; ψ_B and ψ_M are driver disutilities under the baseline contract and under the monitoring contract, respectively (corresponding to $\psi(e,r)$ for respective effort-risk choices); and h is the firing cost the owner incurs in case of a separation. In general, parameters are identified in a nonlinear system of equations in such a way that all identifying moments contribute to the identification of all parameters, at least within the corresponding control or treatment group. Concretely, given empirical moments on the rehiring probability $\mathbb{E}[p(t)|e,r]$, driver contract valuation U, and the target T, $\{\bar{u}, \psi_B, h\}$ are identified in the control group. Similarly, given empirical moments on the rehiring probability $\mathbb{E}[p(t,e,r)|e,r]$ and the target T, ψ_M is (over-)identified in the treatment group.[53] We discuss the identification of these parameters in greater detail below and visualize the identifying relationships in Appendix Figures A.11 to A.13. To further illustrate channels of identification, it is instructive to distinguish between direct and indirect identification of each parameter, which we do in turn.

Outside option $\bar{u}$. This parameter appears directly in the equation for the rehiring probability

$$\mathbb{E}[p(t)|e,r] = 1 - \frac{G(T|e,r)\left(T - \mathbb{E}[y|e,r,y<T]\right)}{\delta U - \bar{u}} \tag{6}$$

as well as the owner's optimal target choice equation via her contract valuation problem:

$$T = \arg\max_{\tilde{T}\in(0,\bar{T})} V(\tilde{T})$$

[53] Note that we have no measure of driver contract valuation U in the treatment group since we measure it only at baseline, before any units are treated. Hence we can no longer use this moment for identification of parameters in the treatment group.

where $\bar{T} = \mathbb{E}[y|e,r] - \bar{u}(1-\delta)/\delta - \psi(e,r)$ and with

$$(1-\delta)V(T) = T - G(T|e,r)\left(T - \mathbb{E}\left[y|e,r,y<T\right]\right)\left(1 + \frac{h}{\delta U - \bar{u}}\right) - \mathbb{E}\left[c|r\right]. \qquad (7)$$

Note, specifically, that an increase in $\bar{u}$ is associated with a decrease in the rehiring probability: if the driver has other attractive options besides the contract with the owner, the resulting contract is harder to sustain. Similarly, an increase in $\bar{u}$ decreases the owner's present-discounted contract valuation $(1-\delta)V(T)$, as she needs to offer the driver larger share of the surplus to satisfy his participation constraint. As a consequence, a higher outside option leads to a lower target T.

In addition, although the outside option does not appear directly in the driver's contract valuation, the direct effects described in the previous paragraph indirectly induce changes in U as well. Specifically, the lower target induced by a higher outside option increases the driver's contract valuation. However, note that we can only observe contract valuation in the control group as we were only able to measure it at baseline, so U only aides identification of $\bar{u}$ indirectly through $\mathbb{E}\left[p(t,e,r)|e,r\right]$ and T in the treatment group.

Driver disutilities ψ_B and ψ_M. Turning to the disutility of contracted driving choices, it is directly identified from the driver's contract valuation equation:

$$(1-\delta)U = \mathbb{E}\left[y|e,r\right] - T - \psi(e,r) \qquad (8)$$

for $\psi(e,r)$ being either ψ_B or ψ_M, depending on whether the driver is in the control group or the treatment group. Higher disutility causes present-discounted valuation to drop, which is the only direct channel of identification. However, in addition to this direct channel, and crucially for disutility in the treatment group ψ_M, where we cannot observe U, the rehiring probability and the target contribute indirectly to the identification of disutility. In particular, higher disutility lowers driver valuation and thereby increases the rehiring probability, see equation (6). Additionally, lower driver valuation decreases the owner's valuation by lowering the target. That is, if driving generates more disutility, then the owner needs to compensate the driver by setting a lower target and thereby increasing his residual salary.

Firing cost h. Finally, the cost incurred by the owner in case of separation is directly identified from the optimal target equation through the owner valuation, i.e. equation (7).

The logic is that a higher firing cost decreases the owner's valuation of the contract in a way that incentivizes her to lower the target so as to decrease the probability of firing. The higher target then indirectly increases driver utility.

Calibrated parameters: discount rate δ and subsistence income w. In addition to the four parameters that are identified in the structural model, we need to specify two calibrated parameters because we do not have any additional linearly independent moments for identification that do not rest entirely on functional form assumptions.[54] These calibrated parameters are the discount rate δ and subsistence income w[55] We choose a number of alternative values for these calibrated parameters to demonstrate the robustness of our conclusions to these choices.

Empirical sensitivity. Andrews et al. (2017) provide a formula to assess the sensitivity of parameter estimates to moments, which is informative about how much each of the moments contributes to the identification of the parameters. This formula is $\Lambda = (\mathbf{J'WJ})^{-1} \mathbf{J'W}$, where $\mathbf{J}$ is the 3×3 Jacobian matrix of derivatives of each of the three moments $\mathbb{E}[p(t)|e,r]$, U, and T with respect to each of the three parameters $\bar{u}$, ψ_B, and h; and $\mathbf{W}$ is a weighting matrix (see below). Columns of Λ show the sensitivity in dollars of a given parameter estimate to a one unit change in each of the moments (i.e. the rows of Λ).

Figure A.10 shows an estimate of Λ in three separate panels, in which each panel represents a column of Λ corresponding to one of the three parameters. Looking at the scales of the three panels, we can see that the three parameters exhibit quite different magnitudes of sensitivity to the identifying moments: $\bar{u}$ is most sensitive, whereas ψ_B is least sensitive. This is intuitive as $\bar{u}$ appears in the denominator of the identifying equations, wheras ψ_B enters linearly in the numerator. Looking across moments, we can see that parameters are quite sensitive to $\mathbb{E}[p(t)|e,r]$ and T, but rather insensitive to U. Given that U is an order of magnitude or several orders of magnitude larger than the other two moments, this also makes sense. We can also see that $\mathbb{E}[p(t)|e,r]$ and T push $\bar{u}$ and h into the same directions, whereas they push ψ_B into the opposite direction.

We provide further evidence of the mechanics of identification and the sensitivity of

[54]There are further moments that are linearly independent, such as owner profit $\mathbb{E}[t|e,r] - \mathbb{E}[c|r]$ and driver salary $\mathbb{E}[y|e,r] - \mathbb{E}[t|e,r]$, but these moments are nonlinear functions of moments already used for identification. Hence, using these moments would amount to identification off of functional form alone.

[55]Evidence from other low-income contexts that suggest a discount factor of 0.99 is reasonable (Anderson et al., 2004). However, the qualitative dynamics of the model remain the same for alternative factors of 0.992 and 0.988. Discount factors beyond these ranges do not converge.

parameter estimates to moments in Appendix Figures A.11 to A.13. In each of these figures, we use the cross-sectional data as described below to generate simulated data for a range of parameter values under the assumption that the model holds (that is, according to the structural moments as described below). Confidence intervals are generated by estimating the standard error of the mean at each parameter value. The figures show that parameters move with structural moments as expected and respond only moderately to changes in parameters, and the inverse function approximates the sensitivity of parameters to changes in the simulated moments. For example, in Figure A.12, we can see that an increase of ψ_B by one dollar is associated with about a one dollar decrease in the target T and a $20 decrease in driver valuation.

E.2 Estimation

We conduct the estimation in two steps. In the baseline estimation step, we seek to estimate $\theta = \{\bar{u}, \psi_B, h\}$, whereas in the estimation under monitoring, we estimate the remaining parameter ψ_M, holding the other parameters fixed. With estimates of the baseline parameter θ, we can estimate the owner's valuation of the contract V, which, together with the driver's valuation U, allows us to estimate the joint surplus generated by the relational contract.

Data Preparation. To prepare the panel data for estimation, it is necessary to create a collapsed, cross-sectional dataset for the $N_c = 129$ control owners and the $N_t = 126$ treatment owners that comprised our experiment. To do so, we first use the full panel data to estimate the standard deviation of revenue (s_i) for each bus. We then collapse the panel to the owner level calculating owner-specific averages for revenue, repair costs, target, salary, and separation probability. We winsorize all variables to remove outliers arising from the unbalanced structure of the panel.

Next, we deal with data problems in the moments for firing probability and the driver's value of the job. Specifically, due to the relatively short duration of the experiment, we observe only a small number of separations with most owners retaining their driver for the duration. This leads to owner specific estimates of $\mathbb{E}[p(t)|e, r]$ to be one, which is incompatible with the estimator. Second, a non-negligible number of drivers reporting zero value of the job due, we believe, to a mis-understanding of the question. To deal with both of these issues, we predict each outcome via OLS using baseline characteristics of the owner, driver, and contract and uses these predicted values for estimation.[56]

[56] Specifically, we use driver education, driver risk-preferences, owner-driver tenure, and the baseline target

Finally, we perform two more data transformations. First, we impose an average difference between the treatment and control samples to match the last three month treatment effects estimated via the reduced form regressions. This is necessary because these estimated treatment effects are not preserved when collapsing the data. Second, we adjust estimated average revenue when this value is inconsistent with the more precisely estimated average target—e.g. if the target is set above the average revenue this would imply the driver has a zero expected salary.

Baseline Estimation. The cross-sectional data for owner-driver pairs is

$$\mathbf{X}_i = (p_i, U_i, T_i, y_i, s_i, c_i)'$$

where $i = 1, ...N$ refers to an owner-driver pair, p_i is rehiring probability, U_i is driver contract value, T_i is the target, y_i is average revenue, s_i is the standard deviation of revenue, and c_i is average repair costs. Let $\widehat{\boldsymbol{\pi}}_i = (p_i, U_i, T_i)$ be the three targeted reduced-form moments (rehiring probability, driver valuation, and target), and $q(\mathbf{X}_i, \theta)$ be the corresponding structural moments predicted by the model. Specifically, the structural moments are given by

$$q(\mathbf{X}_i, \theta) = \begin{bmatrix} q_p(\mathbf{X}_i, \theta) \\ q_U(\mathbf{X}_i, \theta) \\ q_T(\mathbf{X}_i, \theta) \end{bmatrix} = \begin{bmatrix} 1 - \frac{1}{\delta U - \bar{u}} \left\{ G(T) \left(T - \mathbb{E}\left[Y_i | Y_i < T \right] \right) \right\} \\ \frac{1}{1-\delta} \left\{ Y_i - T - \psi_B \right\} \\ \arg\max_T \frac{1}{1-\delta} \left\{ T - G(T) \left(T - \mathbb{E}\left[Y_i | Y_i < T \right] \right) \left(1 + \frac{h}{\delta U - \bar{u}} \right) - c_i \right\} \end{bmatrix}$$

where $G(\cdot)$ is the cdf of a Normal with mean Y_i and standard deviation s_i. $q_p(\mathbf{X}_i, \theta)$ corresponds to the rehiring probability in equation (6); $q_U(\mathbf{X}_i, \theta)$ to driver valuation in (8), and $q_T(\mathbf{X}_i, \theta)$ to the optimal target in the owner valuation problem (7). Specifically, the optimal target is the target that maximizes the owner objective function across a grid of 1,000 potential targets ranging from zero to an upper bound given by $Y_i - \bar{u}(1 - \delta)/\delta - \psi(e, r)$ for any guess of $\{\bar{u}, \psi(e, r)\}$, which guarantees that the driver's participation constraint $\delta U \geq \bar{u}$ holds. Note that the structural moments $q(\mathbf{X}_i, \theta)$ only use data on revenue Y_i, its standard deviation s_i, and the repair costs c_i, whereas driver valuation U and the target T are based on the optimal choices according to model equations (8) and (7) rather than their moments U_i and T_i that we observe in the data.[57]

value.

[57]The fact that our structural moment consists both of data and parameters is the reason why our estimator does not fall into the class of minimum distance estimators: while the reduced-form moment $\widehat{\boldsymbol{\pi}}_i$ consists only of data, our structural moments $q(\mathbf{X}_i, \theta)$ consist of both data and parameters. In contrast, classical minimum

The Generalized Method of Moments (GMM) estimator then minimizes the distance between the reduced-form and structural moments: $m(\mathbf{X}_i, \theta) = \widehat{\boldsymbol{\pi}}_i - q(\mathbf{X}_i, \theta)$. Specifically, letting $M(\mathbf{X}^c, \theta) = N_c^{-1} \sum_{i=1}^{N_c} m(\mathbf{X}_i, \theta)$ be the 3×1 vector of average moments in the control group, where $\mathbf{X}^c = \{\mathbf{X}_i\}_{i=1,\ldots,N_c}$ is the control group data, the estimator is

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} M(\mathbf{X}^c, \theta)' \mathbf{W} M(\mathbf{X}^c, \theta)$$

with weighting matrix $\mathbf{W}$ being the inverse variance-covariance matrix of the estimation moments, i.e. $\mathbf{W} = Var(m(\mathbf{X}_i, \theta))^{-1}$. Θ is a compact set, which we search across with 500 random initial values, selecting the initial value with the lowest objective function value for the reported estimate.

Inference in baseline estimation. Given that we use the optimal weighting scheme $\mathbf{W} = Var(m(\mathbf{X}_i, \theta))^{-1}$, the asymptotic variance of θ is

$$Var(\widehat{\theta}) = \left(\mathbf{J}(\theta)' \mathbf{W} \mathbf{J}(\theta)\right)^{-1}$$

where $\mathbf{J}(\theta) = \nabla_\theta m(\mathbf{X}_i, \theta)$ is the 3×3 Jacobian matrix of moment derivatives with respect to parameters, with each column representing the derivatives of each moment with respect to one parameter. The estimated standard errors of $\widehat{\theta}$ are the square root of the diagonal of $Var(\widehat{\theta})$ after plugging in sample estimates of each object.

To estimate the standard errors of the mean of the reduced-form and structural moments, we proceed as follows. For reduced-form moments $\widehat{\boldsymbol{\pi}}_i$, we use $\sqrt{Var(\widehat{\boldsymbol{\pi}}_i)/N_c}$. For structural moments $q(\mathbf{X}_i, \theta)$, we use the Delta Method:

$$\sqrt{\nabla_\theta q(\mathbf{X}_i, \widehat{\theta})' Var(\widehat{\theta}) \nabla_\theta q(\mathbf{X}_i, \widehat{\theta})}.$$

Estimation under monitoring. Owners who were randomized into the monitoring treatment face the same hiring environment as those who were not, and their drivers face the same labor market conditions as their counterparts working for control owners. As a result, we fix the outside option $\bar{u}$ and the firing cost h at the values estimated under baseline conditions in the control sample. In contrast, the disutility of driving is likely to change as $\psi(e, r)$ moves from the baseline contract effort-risk profile (e_B, r_B) to the monitoring profile (e_M, r_M) with reduced risk. Consequently, the goal in the second estimation step is to

distance (CMD) estimators additively separate data and parameters as in $\widehat{\boldsymbol{\pi}}_i - q(\theta)$.

estimate $\psi_M = \psi(e_M, r_M)$, holding $\bar{u}$ and h fixed.

We only have a measure of driver contract valuation U_i at baseline. Therefore, we only have two reduced-form moments, p_i and T_i, corresponding to the structural moments $\mathbb{E}\left[p(t,e,r)|e,r\right]$ and T for the treatment group. The estimator under monitoring is thus over-identified. Let

$$m_M(\mathbf{X}_i, \psi; \widehat{\theta}) = \begin{bmatrix} p_i - q_p(\mathbf{X}_i, \widehat{\theta}) \\ T_i - q_T(\mathbf{X}_i, \widehat{\theta}) \end{bmatrix}$$

and $U = \frac{1}{1-\delta}\left\{Y_i - T - \psi_M\right\}$ in the structural moments $q_p(\mathbf{X}_i, \widehat{\theta})$ and $q_T(\mathbf{X}_i, \widehat{\theta})$. Defining $M_M(\mathbf{X}^t, \theta) = N_t^{-1}\sum_{i=1}^{N_t} m(\mathbf{X}_i, \theta)$ as the 2×1 vector of average moments in the treatment group, where $\mathbf{X}^t = \{\mathbf{X}_i\}_{i=1,\ldots,N_t}$ is the control group data, the second-step GMM estimator is then

$$\widehat{\psi}_M = \arg\min_{\psi \in \mathbb{R}} M_M(\mathbf{X}^t, \psi; \widehat{\theta})' \mathbf{W}_M M_M(\mathbf{X}^t, \psi; \widehat{\theta})$$

with weighting matrix $\mathbf{W}_M = Var\left(m_M(\mathbf{X}_i, \psi; \widehat{\theta})\right)^{-1}$. We use the estimated disutility from the first step, $\widehat{\psi}_B$, as the initial value.

Inference under monitoring. Since we use the estimates from the first step in the second step, we use a bootstrap with 500 iterations wrapped around both estimation stages to estimate the standard error of $\widehat{\psi}_M$.

Validation. We can examine to what extent the model captures variation in the data well by comparing untargeted reduced-form and structural moments for which we have data but are not included in the estimation because they are not direct functions of the parameters. Specifically, we inspect driver salary defined as the difference between the expected renenue and the expected transfer (i.e. the residual claim), $\mathbb{E}\left[y|e,r\right] - \mathbb{E}[t(y)|e,r]$, as well as the owner profit defined as the difference between expected transfer and expected costs, $\mathbb{E}[t(y)|e,r] - \mathbb{E}\left[c|r\right]$, where

$$\mathbb{E}[t(y)|e,r] = T - G(T|e,r)\left(T - \mathbb{E}\left[y|e,r,y<T\right]\right).$$

These untargeted moments are shown at the bottom of Panel C in Tables 6 and 7. In each case, both for the baseline estimation as well as the treatment effect (i.e. monitoring) estimation, the difference between reduced-form and structural moments is small and not statistically significant, suggesting that our estimator performs well on untargeted moments as well.