American Economic Review 2024, 114(10): 3119–3160
https://doi.org/10.1257/aer.20210987

Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit†

By Erin M. Kelley, Gregory Lane, and David Schönholzer*

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 to utilize monitoring technologies to overcome problems of moral hazard and enhance their profitability. (*JEL* D22, D24, D82, J41, L25, L92, O14)

Small and medium-sized firms account for the majority of businesses in low-income countries, they employ over half of the population, and they account for more than 40 percent 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). Any firm seeking to expand needs to grapple with the challenges associated with managing their workforce. 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 (Lucas 1978; Shahe Emran, Morshed, and Stiglitz 2021; Akcigit, Alp, and Peters 2021). 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 forthcoming), there is

* Kelley: Harris School of Public Policy, University of Chicago (email: erinmkelley@uchicago.edu); Lane: Harris School of Public Policy, University of Chicago (email: laneg@uchicago.edu); Schönholzer: Department of Economics, UC Santa Cruz (email: dschoenh@ucsc.edu). Rema Hanna was the coeditor for this article. 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). AEA registration number is AEARCTR-0001482. 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.
† Go to https://doi.org/10.1257/aer.20210987 to visit the article page for additional materials and author disclosure statement(s).

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 the contracts they offer, boost profits, and stimulate growth. 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 business outcomes to their actions, problems of moral hazard can persist for some time and limit firm performance.

In this context, monitoring technologies can enhance outcomes by granting owners increased visibility into their employees' operations. 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 manage a single minibus. 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 continue 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.

Interpreting the impacts of monitoring technologies requires 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, and he may not be rehired if he persistently fails to meet the target. The owner is liable for major expenses accrued during the day. This *target contract* can be found in other parts of the world as well (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.[1] In other words, we show 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 high effort *and* excessive risk taking. 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

[1] 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 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.

contract space by making effort and risk observable, 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 use the system to monitor their drivers' 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 drivers' daily revenue target by 4.9 percent (p-value $=$ 0.114), and driving behavior is geared toward more effort and less risk taking. Treated drivers increase the number of hours they spend on the road by 9.8 percent (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 (and salary) as before. This lowers repair costs by 44.6 percent (p-value $=$ 0.037) and contributes to substantial increases in daily profit for the owners. In month 4 of the experiment, profits increase by 13.7 percent (p-value $=$ 0.046). 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 (10.5 percent) more likely to own an additional vehicle than control owners by the end of the study (p-value $=$ 0.091).

As these technologies become 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 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 \$507, and the owner at \$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 \$83 (4 percent) 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 \$20 (4 percent) under monitoring. This is primarily because they incur greater disutility from having to drive in a less risky way. The impact of monitoring on total welfare is therefore small, albeit imprecisely estimated. 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 percent of drivers say they preferred driving with the device because it improved trust with the owner, and owners devote 30 percent more to drivers in a trust game at endline.[2]

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

[2] Our welfare evaluation also abstracts from welfare effects on passengers—an important consideration in light of how dangerous minibuses are. Monitoring did not significantly affect the number of accidents.

studying the impact of monitoring technologies in a specific environment where owners have one worker they interact repeatedly with and whose behavior they cannot always easily observe. If 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. Moreover, if repeated interactions make it easier to attribute poor business outcomes to an employee's performance, the scope for monitoring may be reduced.

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 low-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 forthcoming). Finally, monitoring capacity remains 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. Our paper most closely resembles work on managerial deficits, which primarily 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, Bjorvatn, and Tungodden 2015; McKenzie and Woodruff 2017). Yet a few papers provide evidence that managing employees is a challenge in low-income countries. They show that many small businesses do not trust their workers, which discourages hiring (Caria and Falco forthcoming); others use complex rental arrangements to avoid merging and managing a larger labor force (Bassi et al. 2022); and some struggle to adopt new technologies because of employees' misaligned incentives (Atkin, Khandelwal, and Osman 2017). Our paper focuses directly on employee management, documenting how the provision of information to the firm about employee behavior affects informal contracts and firm profitability.

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 (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 (online Appendix Figure A.1). 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 US trucking industry. We build on this work by introducing exogenous variation in the usage of monitoring technologies through an RCT and capturing high-frequency data on contracts and worker behavior in a low-income country context, where relational contracts are more prevalent and monitoring devices reduce rather than eliminate information asymmetries. 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. A number of papers find that monitoring systems improve labor productivity (Gosnell, List, and Metcalfe 2020) and reduce moral hazard (Liu, Brynjolfsson, and Dowlatabadi 2021; Gertler et al. 2023). Prior research suggests that monitoring systems may affect firms differently in low-income countries. The quality of management practices is typically lower (Bloom et al. 2013), which could prevent firms from harnessing the benefits of monitoring technologies. Moreover, contract enforcement is weak and workers are poor, which could also limit 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 constraint to firm profits in low-income countries.

Finally, we contribute to a growing literature documenting the impact of policies that improve the efficiency of transportation networks within cities (Hanna, Kreindler, and Olken 2017; Kreindler 2020; Tsivanidis 2019). In a companion paper we study how the introduction of monitoring affects the safety of informal transit systems (Lane, Schönholzer, and Kelley 2022), which builds on work by Habyarimana and Jack (2011, 2015).

The rest of the paper is organized as follows. In Section I we present relevant context about the industry. In Sections II and III, we describe the experimental design and the data we collect. Section IV develops a theory of contracting in this industry. We present reduced-form results of the experiment in Section V. Section VI provides results from our structural estimation, and Section VII contextualizes its welfare implications. Section VIII concludes.

I. Context

A. *Minibus Industry*

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 services are primarily operated by private entrepreneurs who own small fleets of minibuses, commonly referred to as *matatus*. Rough estimates suggest that approximately 15,000 matatus operate within the city, supporting a massive industry that contributes up to 5 percent of the country's GDP (Kenya Roads Board 2007). Matatu owners typically purchase 14-seat minibuses. They obtain licenses for operating on specific routes and manage vehicle operations themselves. Route management is overseen by Savings and Credit Cooperatives (SACCOs), which coordinate centralized activities such as ensuring adherence to regulations set by the National Transport and Safety Authority. Passengers board matatus at various points along their route and pay their fare in cash (Bruun and Behrens 2014).

Minibus owners in Kenya (and other countries) hire drivers on "target" contracts: the owner hires a single driver to operate their vehicle and sets a daily revenue target for the driver. The driver is the residual claimant.[3] If the driver misses the

[3] In Kenya, the driver is accompanied by a fare collector whom 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.

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 (Behrens, McCormick, and Mfinanga 2015). 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.

The industry is widely perceived to suffer from several inefficiencies (McCormick et al. 2013; Behrens, McCormick, and Mfinanga 2015; Mutongi 2017). A lack of enforcement creates incentives for drivers to operate on unlicensed routes, where they pay substantial fines when they are caught. 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 Report on Road Safety, approximately 3,000–13,000 people die annually from traffic incidents in Kenya, and at least 62 percent of cases involve matatus (Odero, Khayesi, and Heda 2003; WHO 2013).

B. *Moral Hazard*

To document the extent of moral hazard in this environment, we conducted descriptive surveys 5 years after the RCT (in 2022) with 150 matatu owners operating across Nairobi. We provide further details about this exercise in online Appendix 2. We find that matatu owners are unable to observe their drivers' behavior with complete accuracy: approximately 65 percent of owners state that they can only sometimes determine if their driver is driving recklessly, while 32 percent report that they can rarely tell. Similarly, 80 percent of owners claim that they can only sometimes determine if their driver is responsible for damage to the vehicle, with 20 percent saying that they can rarely tell.[4] Lastly, 56 percent of owners report that they can only sometimes tell if their driver is operating off route, while 36 percent 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 percent of owners reported that they sometimes or always faced challenges managing their drivers, while 25 percent said that they rarely or never faced them. 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 percent of the owners agreed that 50 percent 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 percent of owners

[4] 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 severe accidents.

believe their own drivers can improve their performance in some way. Hiring a new driver who will outperform an existing one is therefore difficult and risky because the quality of drivers is fairly low on average. This means that owners may retain drivers even if they are not entirely satisfied with their on-the-job performance, and problems of moral hazard could persist for some time.

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 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 report these results in online Appendix 2.

C. *Monitoring*

GPS tracking devices began entering the Kenyan market around 15 years ago. Many insurance providers also began mandating that minibuses install GPS trackers for security reasons (Business Daily Africa 2009). While minibus owners (and drivers) were familiar with tracking systems, most 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 (described in Section II).

Our descriptive surveys from 2022 demonstrate that matatu owners perceive the importance of monitoring technologies and find value in using them. Approximately 50 percent of owners say that GPS tracking devices are always useful, while 40 percent of owners find them sometimes useful. Furthermore, 50 percent 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 traveled, and monitoring instances of speeding, sharply braking, and sharply turning.

Finally, we also compile evidence to show that monitoring technologies can be useful in settings beyond the one we study. We conduct surveys with other businesses in Kenya and compile anecdotal and empirical evidence from high-income countries, all of which suggest there is value in monitoring technologies. We present this evidence in online Appendix 2.

II. Experimental Design

A. *Tracking Device and Software*

To understand the impact of monitoring on the matatu industry, we developed the *SmartMatatu* monitoring system with a Kenyan technology company (Echo Mobile). We developed our own system because available alternatives on the market were either too costly or not sophisticated enough. The R&D process lasted more than one year and benefited from extensive discussions with matatu owners. The physical tracking units were procured from a company in the United States (CalAmp). The tracking device had a GPS and gyroscope, which captured the vehicle's location and its vertical/lateral/forward and backward acceleration at 30-second intervals. The device relied on GPRS to send the information from the tracker to our servers via the cell phone network. The data were further processed on the server to provide daily measures of the vehicle's mileage, the number of hours the ignition was on, average and maximal speed, and the number of speeding, overacceleration, sharp braking, and sharp turning alerts. Finally, an API call was generated each time the owner used the app to request data from the server.

We designed a novel mobile application to convey information to owners in a user-friendly way (Figure 1). The app's first tab was a map of Nairobi and presented the vehicle's real-time location. By entering a specific time interval into the phone, the app could display the exact routes the matatu traveled over this time period. This first tab conveyed a more accurate measure of costly driving because owners could see if the driver was operating on roads that were known to damage vehicles. The second tab displayed the safety alerts captured by the device. The final tab conveyed a summary of the driver's effort and safety. The effort section listed the total mileage covered and the duration the ignition was on that day. Finally, the SmartMatatu app was designed to collect daily information from owners about the business.

The technology cost approximately \$125, which reflects the cost of the device itself (\$85), installing it (\$25), and storing/processing the data (\$15). Two software engineers maintained the app for the duration of our study at the cost of about \$83 per unit. Hardware maintenance and replacement costs were negligible, as there were fewer than five devices that needed to be replaced throughout the course of our study. The cost of providing our device is comparable to existing business training programs. Van Lieshout and Mehtha (2017) report the average cost for offering a week-long business training course in 18 different countries to be about \$177.[5]

B. *Treatment Assignment*

In 2015, we contacted SACCOs operating across nine major commuter routes in Nairobi and organized meetings with matatu owners to present the study's goals and methodology. We registered interested owners who satisfied 3 conditions: they owned a single 14-seater matatu, they managed it themselves, and they employed a driver rather than driving the minibus themselves. We informed all owners that we would be placing a monitoring device in their vehicle and they would be required to provide daily information about their business. We also mentioned that a random subset of owners would be selected to receive information immediately, while others would have to wait six months before gaining access to the information for a two-month period. It took four months to recruit enough participants across the nine routes (online Appendix

[5] We adopt the prevailing approach to pricing business development services, which involves considering pricing at or above *marginal cost* (Karlan and Valdivia 2011; Drexler, Fischer, and Schoar 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.

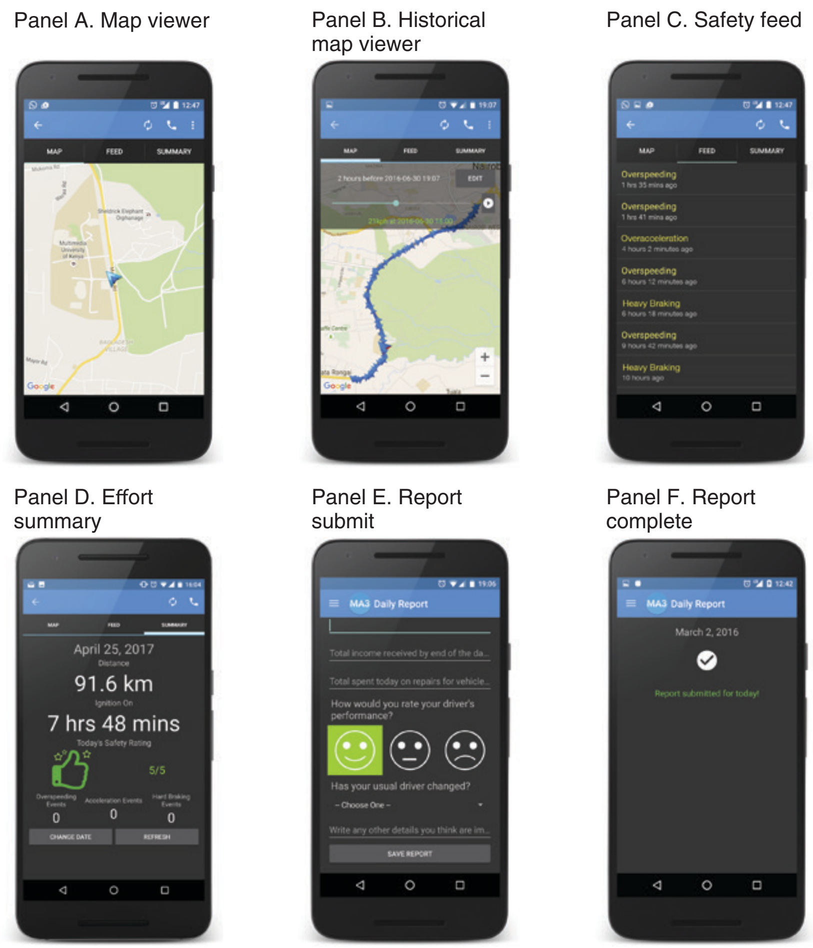

FIGURE 1. MOBILE APP "SMARTMATATU"

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 A.2). We registered 255 owners, whom we randomized into treatment (126) and control (129), stratified by route.[6] Spillover concerns are minimized, as SACCOs typically have hundreds of members, and our intervention would have impacted less than 3 percent of each route.

We conducted installations and trainings from November 2016 to April 2017 (online Appendix Figure A.9). The field team scheduled a time to meet each owner

[6]The recruitment process took time because of the difficulties we faced scheduling meetings with owners who had many other commitments. While some owners remained hesitant to participate, this was not for a lack of enthusiasm but rather a reluctance to invest time and money into learning a new technology.

individually at a location of their choosing. Every owner was compensated for the time their vehicle spent off-road to perform the installation of the device with a onetime payment of KSh 5,000 ($50). For both the treatment and control groups, we installed trackers under the vehicle's dashboard to prevent tampering and provided owners with an Android smartphone with our SmartMatatu app pre-installed. The app only provided tracking information to owners randomized into the treatment group, who received an additional 30 minutes of training on how to navigate this information. 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 six-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 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. This ensured that drivers could only learn about the specific data collected by the device from the owner (though drivers likely correctly inferred from the device's location that it could not monitor exact revenue or repair costs). 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 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.[7]

III. Data and Descriptive Statistics

A. *Data Collection*

We first administered a baseline survey during the tracker installations. For owners, we collected basic demographics, employment history, features of the matatu, and their relationship with the current driver. Similarly, for drivers, we asked about driver demographics, experience as a driver, and their relationship with the current owner. We captured the driver's stated value of the contract by asking them to consider how much they would have to be paid to give up the job. We also used games to gauge drivers' risk aversion and drivers'/owners' propensity to trust one another (Sprenger 2015). To measure risk, we asked respondents whether they preferred to receive KSh 500 ($5) for certain or play a lottery to win KSh 1,500. The trust game was similar to Berg, Dickhaut, and McCabe (1995): we presented owners with KSh 500 and asked them to select an amount to be placed back in an envelope. This amount was then tripled and delivered to some driver (other than their own) who

[7] We also cross-randomized across treatment and control a driver cash incentive treatment beginning in month 5 of the experiment to encourage safer driving. This is explored in a separate paper.

decided how much to keep for himself and how much to return to the owner. The amount the owner chose to place in the envelope was recorded in the survey. At the end of the six-month period, we also conducted an endline survey focused on business investment decisions. Finally, we ran a willingness-to-pay experiment, offering owners two additional months of monitoring information through the app.

Next, we collected daily data from owners and drivers. For owners, we relied on our SmartMatatu app. Owners were reminded daily via push notification to submit data through the app, including 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. Owners received KSh 40 via M-Pesa (a mobile money service) for each submission.

We collected a distinct set of outcomes from drivers through SMS surveys (because the drivers did not receive smartphones). Specifically, we asked about whether the vehicle was on the road, the amount of revenue they collected, and the residual revenue they kept as a salary. We emphasized that all of the data they shared was confidential and would not be shared with the owner. Drivers were compensated KSh 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 (online Appendix Table A.1). We regress revenue and salary on an objective measure captured by the tracker (number of miles traveled), an indicator for treatment, and the interaction 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/salary between the treatment and control groups. This reduces concerns about differential reporting.

Finally, we rely on the tracking device data. We use 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. This provides a measure of how far the driver is deviating from the actual route. Online Appendix 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.

B. *Descriptive Statistics*

Owners in our sample are predominantly self-employed men in their late thirties (Table 1). They have managed their own vehicles for four years and possess eight years of experience in the industry on average. Drivers in our sample are exclusively male, slightly younger than the owners, and have lower levels of education. They have eight years of experience working as matatu drivers on average. The matatu vehicles are primarily imported Japanese minibuses that have been used for approximately 13 years (online Appendix Figure A.4). Some of these vehicles come equipped with special features such as free Wi-Fi, sound systems, or TVs. The average purchase price for these matatus is approximately $6,675.

The owner sets a daily revenue target at baseline of approximately $31 (KSh 3,130), and they report receiving $26 (KSh 2,600) on average from the driver. The target amount set by owners exhibits some variability (standard deviation of

Table 1—Summary Statistics for Owners, Drivers, and Matatus

	All				Treatment mean	Control mean	p-value difference
	Mean	SD	Min	Max			
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 ($)	6,675	2,849	1,800	30,000	6,396	6,947	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 (columns 1, 2, 3, and 4); means for the treatment and control groups (columns 5 and 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 hundreds of Kenyan shillings (KSh , 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.

KSh 446), indicating that owners have some discretion in determining the desired target within industry norms. Variation in the baseline target is most highly associated with differences in the quality of the matatu itself, including the matatu's age, the number of features it has, and its price (online Appendix Table A.2, columns 3 and 4). We do not see evidence that the baseline target is associated with owner-driver tenure or other driver characteristics (online Appendix Table A.2, columns 1 and 2). While we might expect owners who have worked with their drivers for longer to trust them more and assign lower targets, we find no evidence for this.

Drivers report collecting approximately $71 (KSh 7,126) in passenger fares (revenue) throughout the day. They retain approximately $9.07 (KSh 907) as their salary (Table 3) and spend the rest on fuel costs and bribes to the police. 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 (KSh 483). We investigate whether owners can leverage their repeated relationship with drivers to learn how they operate and enforce better outcomes. Online Appendix Table A.3 examines the relationship between four key business outcomes (hours ignition was on, repairs, revenue, and profits) and (i) owner-driver tenure, (ii) driver characteristics, and (iii) driver risk aversion, all 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. Since risk aversion is not observable to the owner, monitoring mechanisms could prove valuable in increasing firms' visibility into drivers' working styles.

Reckless driving is widespread within our sample (online Appendix Figure A.5). We capture reckless driving in three ways: the share of days drivers exceed 75 km/h (speed limit is 50 km/h), the share of days drivers are flagged for sharp braking, and the number of hours drivers deviate from the designated route per day. 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 percent of drivers exceed 75 km/h on more than 20 percent of the days they drive, and approximately 25 percent of drivers consistently exceed this speed limit over half the days they operate the minibus. Next, we find that approximately 50 percent of drivers break sharply on 20 percent of the days they operate the vehicle, with around 25–30 percent of drivers braking sharply over half the days they operate the minibus. Finally, we observe that 50 percent of drivers spend approximately 3 hours per day traveling 400 meters 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 percent of drivers spend an additional 1.5 hours beyond their designated route throughout the day.

There is some turnover between owners and drivers. 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. The average employment tenure of 14 months in Table 1 is heavily influenced by a small number of long-lasting relationships. The likelihood of an owner and driver separating (either through firing or quitting) in our sample is estimated to be approximately 0.1 percent per day. This implies that there is a 99.9 percent chance of the driver being rehired the following day, and the annual probability of driver-owner separation is $1 - 0.999^{365} = 31$ percent. Being fired does not appear to impose substantial reputational costs on drivers. In the descriptive survey we conduct a few years later, over 80 percent of owners acknowledge that dismissed drivers can secure employment with another firm within the SACCO. Owners attribute this to a combination of factors: varying preferences among owners (70 percent), the necessity to settle for available options due to high demand and the scarcity of good drivers (70 percent), and the lack of information provided by drivers regarding if and why they were terminated (60 percent).

IV. Model

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. 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 contract theory. 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 first set 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). Next, we show how the contract changes when monitoring is introduced, which allows the owner to observe some driver choices. We refer to the principal as the female owner and the agent as the male driver.

A. *Setup*

A risk-neutral owner and risk-neutral driver 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 call "risk" and denote by r.[8] 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 $G(\cdot \mid e, r)$, 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 $F(\cdot \mid r)$. Repair costs depend on risk but not effort and accrue entirely to the owner. Conditional on effort and risk, the revenue and cost distributions are independent.[9]

[8] The use of "risk" is nonstandard in the literature (e.g., Ghatak and Pandey 2000). In our model, risk is a second effort dimension with an additional cost to the principal instead of having a mean-preserving effect on the variance of output. We call this "risk" because it corresponds to actions such as risky maneuvers that damage the vehicle or taking an unlicensed route, risking a fine. Despite the agent having multiple choice dimensions, our model does not fall into the class of multitasking frameworks (Holmström and Milgrom 1991) because our monitoring technology does not shift the *relative* observability of the two dimensions.

[9] See online Appendix 5.2 for functional forms assumptions of the technology and preferences.

The owner chooses whether to rehire the driver for the next day with probability $p(\cdot)$—the rehiring schedule. In the baseline environment without monitoring, this rehiring schedule depends only on the transfer $t \in [0, \bar{y}]$.[10] If the owner has access to monitoring, 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.[11] 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.[12]

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 revenue y as well as repair cost c. The driver then transfers t according to his transfer schedule $t(y)$ 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 the case of monitoring. If he is rehired, the game repeats the following day.

B. *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:

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

subject to

$$\text{(i)} \quad U - \bar{u} = \mathbb{E}\big[y - t(y) + p\big(t(y)\big)(\delta U - \bar{u}) | e, r\big] - \psi(e, r) \geq 0$$

$$\text{(ii)} \quad (e, r) \in \underset{(\tilde{e},\tilde{r}) \in \mathcal{S}}{\operatorname{argmax}}\, \mathbb{E}\big[y - t(y) + p\big(t(y)\big)(\delta U - \bar{u}) | \tilde{e}, \tilde{r}\big] - \psi(\tilde{e}, \tilde{r})$$

$$\text{(iii)} \quad t(y) \leq y$$

$$\text{(iv)} \quad y - t(y) + p\big(t(y)\big)\delta U \geq y$$

$$\text{(v)} \quad t(y) \in \underset{\tilde{t} \geq 0}{\operatorname{argmax}}\; y - \tilde{t} + p(\tilde{t})(\delta U - \bar{u}).$$

While the driver ultimately chooses effort and risk, the owner treats them as choice variables for the purpose of designing the contract. The first constraint is the participation constraint, which restricts driver utility to be at least as great as his outside

[10] We consider the possibility that owners may use information on driver risk (through repair costs) in online Appendix 5.4. Yet qualitative and anecdotal evidence supports our modeling choice that repair costs are an unreliable indicator of driver behavior and are thus not included in the rehiring schedule.

[11] This is an important feature of monitoring in informal transit. Even if the owner knows the number of trips a driver took, determining the exact revenue is impossible because they lack information on the number of passengers. Drivers said they were more comfortable with technologies that reveal their choices of effort and risk, compared to ones that allow owners to observe revenue, such as electronic payment systems.

[12] We focus on moral hazard in stable owner-driver relationships, and we do not study adverse driver selection; less than 15 percent of owners get a new driver, limiting the impact of adverse selection in our study.

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 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.[13]

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 up front. 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 the daily target T, which defines the level of transfers above which reemployment is guaranteed. Under the assumption that the owner prefers less risk than the driver, which we discuss more in online Appendix 5.2, we can solve for the optimal transfer and rehiring schedules:

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

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

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 reemployment 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

[13] 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 strategy into account when considering the contract. While it may be possible to incorporate this 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 transferring nothing to the owner.

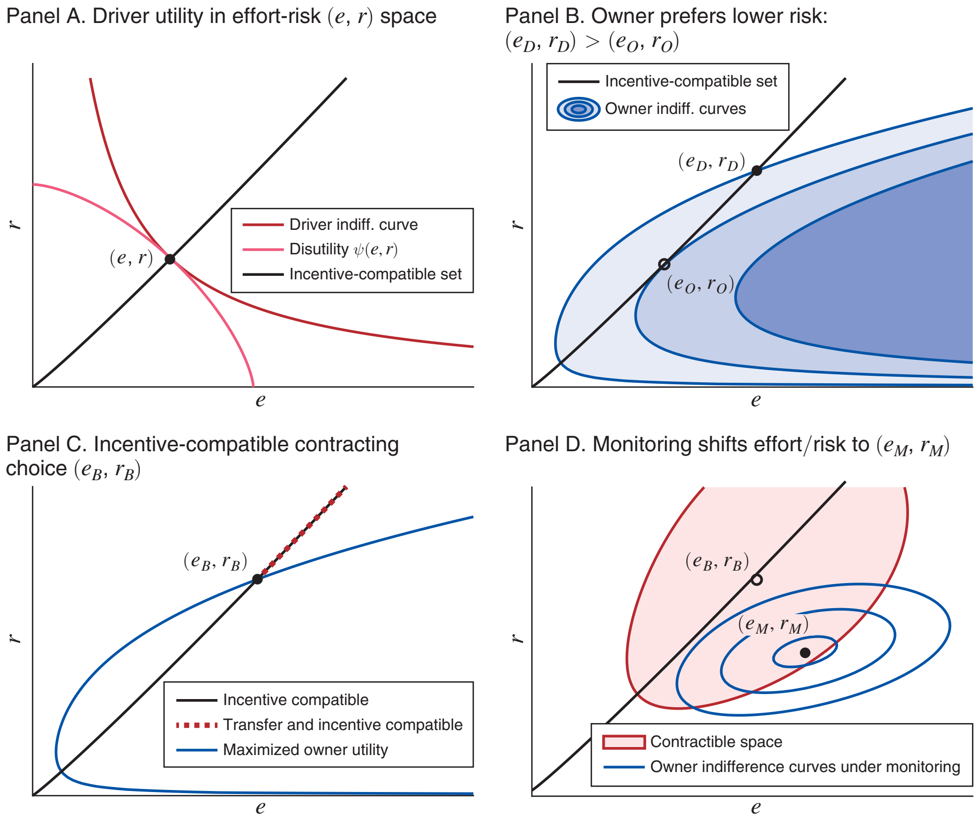

Figure 2. Baseline and Monitoring Contract Intuition

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 in the online Appendix). 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 trade-off in effort and the target.

effort-risk bundles on the driver's incentive-compatible set (Figure 2, panel A). 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 (e_B, r_B) because it is the choice of effort and risk the driver would make if they 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 (Figure 2, panel B), 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. This is her preferred bundle among those that are both incentive compatible and transfer compatible (Figure 2, panel C).[14]

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 online Appendix 5.3, 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.

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.

C. *Introducing Monitoring*

With monitoring, the owner conditions her rehiring schedule on the driver's choice of effort and risk, which she now observes, as well as the transfer: $p(t,e,r)$. The solution to rehiring schedule becomes

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

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 (proofs are in online Appendix 5.3).

(i) **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.

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

(ii) **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.

(iii) **Profits Increase:** Profits will unambiguously increase due to lower repair costs.

(iv) **Targets Fall 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.[15]

(v) **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.[16]

V. Experimental Results

A. *Empirical Contract Characteristics*

Our data show 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 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.

[15] There are two forces that influence the owner's decision to reoptimize 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.

[16] 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 $c(r) \gg c(r_m)$, but $\psi(e, r)$ is only slightly 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.

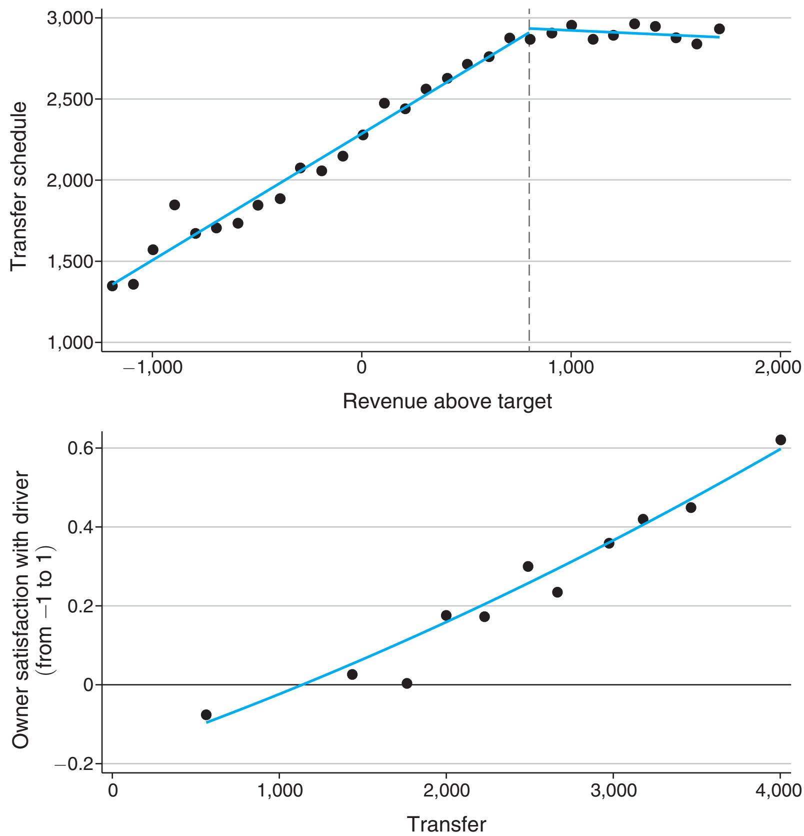

Figure 3. Estimated Transfer Schedule and Owner Satisfaction in Response to Transfer

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 in online Appendix 5.3. The slope extends beyond the target because of subsistence income. Bottom panel: Owner satisfaction with driver as a function of the transfer. Owner satisfaction rises substantially with the transfer, as suggested by $p(t) = 1 - \frac{T - t}{\delta U - \bar{u}}$.

B. *Treatment Take-Up*

We monitor treated owners' usage of the device by tracking the API calls that are generated every time the owner logs into the app and requests different pieces of information. Online Appendix Figure A.6 shows that 70 percent of owners consult the app weekly (panel A), while 50 percent use it daily (panel D). Panel C shows that owners are significantly engaged with the app, using it for eight hours per week by month 6.

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). While these survey responses may be subject to experimenter demand effects, we find that owners in the treatment group are 27 percentage points more likely to say they know about the number of kilometers driven and 45 percentage points more likely to know about the instances

Table 2—Treatment Effects on Reported Knowledge and Monitoring Behavior

	Know mileage (1)	Know off-route (2)	Know revenue (3)	Difficulty monitor (4)	Monitoring time (5)
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 fixed effects	X	X	X	X	X
Matatu observations	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's 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 knows 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.

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. Having access to the information reduces the reported difficulty level of monitoring by just under 2 out of 5 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 see that 72 percent of treated owners report a decrease in the time they spend monitoring (Table 2, column 5). Finally, while we do not have a way to track whether owners follow up with their drivers using the app's information, our conversations with owners and drivers suggest such interactions occurred.

C. *Treatment Effects*

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; and ε_{id} is an error term. We cluster the standard errors at the matatu/owner/driver level.[17] This design allows us to examine treatment effects as they evolve over the six months of the study. This is important because it took a few months for owners to become comfortable with all the features of the monitoring app. We consider the impacts presented below as intent-to-treat (ITT) estimates, as some owners did not consistently use the app.[18]

Our trial was registered on the AEA RCT Registry (# AEARCTR-0001482). The major deviations from this document include only reporting productivity outcomes, not implementing 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.[19]

Driver Choices (*Effort and Risk*).—Monitoring devices allow owners to observe drivers' effort and risk and guide them toward more favorable choices, specifically, encouraging higher effort and lower risk. We proxy driver effort by the number of hours the matatu is operating. We find that operating hours rise steadily until the end of the study (Figure 4 and Table 3). By month 6, treatment drivers spend 1.45 (9.8 percent) more hours on the road per day (p-value $=$ 0.055). This is a significant increase in a context where drivers already work 14-hour days. We also see the number of kilometers increase by 13 kilometers per day on average (13 percent) by month 6 (p-value $=$ 0.062), which corresponds to an extra trip to the city center. While this increase in effort benefits the firm, we may worry about safety externalities for passengers exposed to drivers in their fifteenth hour. However, there is no significant increase in safety-related outcomes.

Next, we analyze changes in risk-taking behavior, particularly the tendency to operate on unlicensed routes. While these routes offer a way to bypass traffic, they are often less well maintained. Off-road driving appeals to drivers, as it allows them to avoid traffic jams, and owners have to cover any associated repairs or fines. Figure 5 shows that treatment drivers are approximately 400 meters closer to the licensed route than control drivers throughout the study (p-value $=$ 0.022). We use the acceleration data to explore whether treated drivers are staying on designated routes that are less bumpy and damaging to the vehicle. We find that the distributions of lateral and vertical acceleration in the treatment group tighten around zero, consistent with a reduction in damaging driving. We can reject equality of treatment and control distributions by applying a K-S test (p-value $=$ 0.001). These results suggest that owners use the monitoring technology to ensure that drivers stay on designated routes that are better maintained and less bumpy.

[17] For precision, we include as controls basic characteristics of the matatu (age, special features) and basic owner demographics (age, sex, employment, experience, Raven's score). We do not control for driver characteristics because drivers can change over the course of the experiment. Our results are qualitatively similar if we add additional matatu/owner controls.

[18] Table 3 presents all the treatment effects by month in table form and corresponds one-for-one to the figures presented below. Online Appendix Table A.4, panel 1 presents the specification detailed in the AEA registry, which pools all months of our data together. Online Appendix Table A.4, panel 2 presents treatment effects from a pooled regression of the same outcomes on indicators for being treated in the first three months and the last three months of the study and provides another way to showcase the different stages of learning.

[19] More broadly, the approach in the paper has been to focus our primary analysis on the most 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.

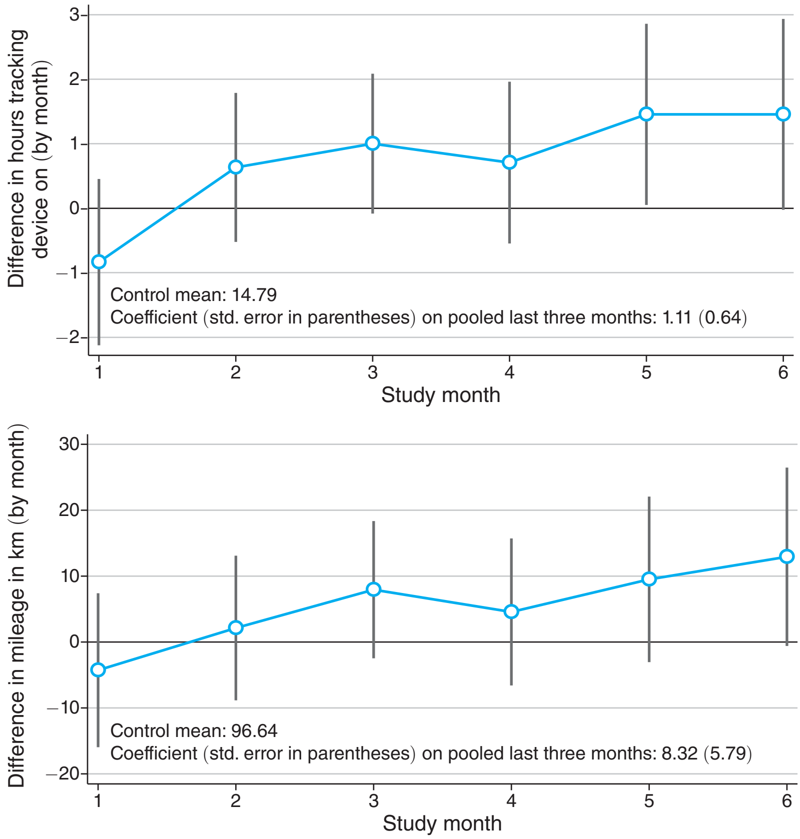

Figure 4. Treatment Effects on Effort

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 percent 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)).

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 month 3, daily repair costs for treatment owners fall by KSh 124 ($1.24) ($p$-value $=$ 0.12). By month 6, daily repair costs are KSh 216 ($2.16) per day lower for treatment owners (p-value $=$ 0.037). These reductions represent a 44.6 percent decrease in daily repair costs. As repair costs constitute a significant expense for owners, monitoring technologies can have a substantial impact on the business.

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 their incentive to do so because they are more likely to be caught. This is unlikely to be the case for larger repairs because they are incurred directly by the owner and will be validated with the mechanic. We create an indicator for whether repair costs exceed KSh 1,000 ($10—eightieth percentile). The probability of incurring a large repair cost decreases significantly in

Table 3—Treatment Effects on Effort, Costs, Revenue, Target, Profits, Salary

	Device on (hours)	Mileage (kilometers)	Repair costs	Repair costs (large)	Revenue	Target	Met target	Gross profit	Salary per hour
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
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	7,126.9	3,057.4	0.43	3,260.5	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 fixed effects	X	X	X	X	X	X	X	X	X
Route fixed effects	X	X	X	X	X	X	X	X	X
Matatu-day observations	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 (all monetary values in KSh). "Repair costs (large)": Owner-reported daily repair costs that exceed \$10. "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's 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.

month 6 (7.71 percentage points, p-value $= 0.042$) (online Appendix Figure A.7 and Table 3, column 4). This suggests that the decrease in the repair costs that we observe cannot be entirely driven by inflated repair costs.

As effort increases while risk taking falls, the impact of monitoring on revenue is ambiguous and depends on which effect dominates. Owners may be willing to accept lower revenue if the reduction in repair costs from less risk taking more than offsets the reduction in expected transfers from lower 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.

Owner Choice (*Target*).—As the revenue collected largely stays the same, we anticipate the target should decrease, as owners need to compensate the driver for the larger disutility from work. 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 6, the daily target amount is KSh 149 (\$1.49) below the control group, representing a 4.9 percent decrease. 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 reoptimize the terms of their

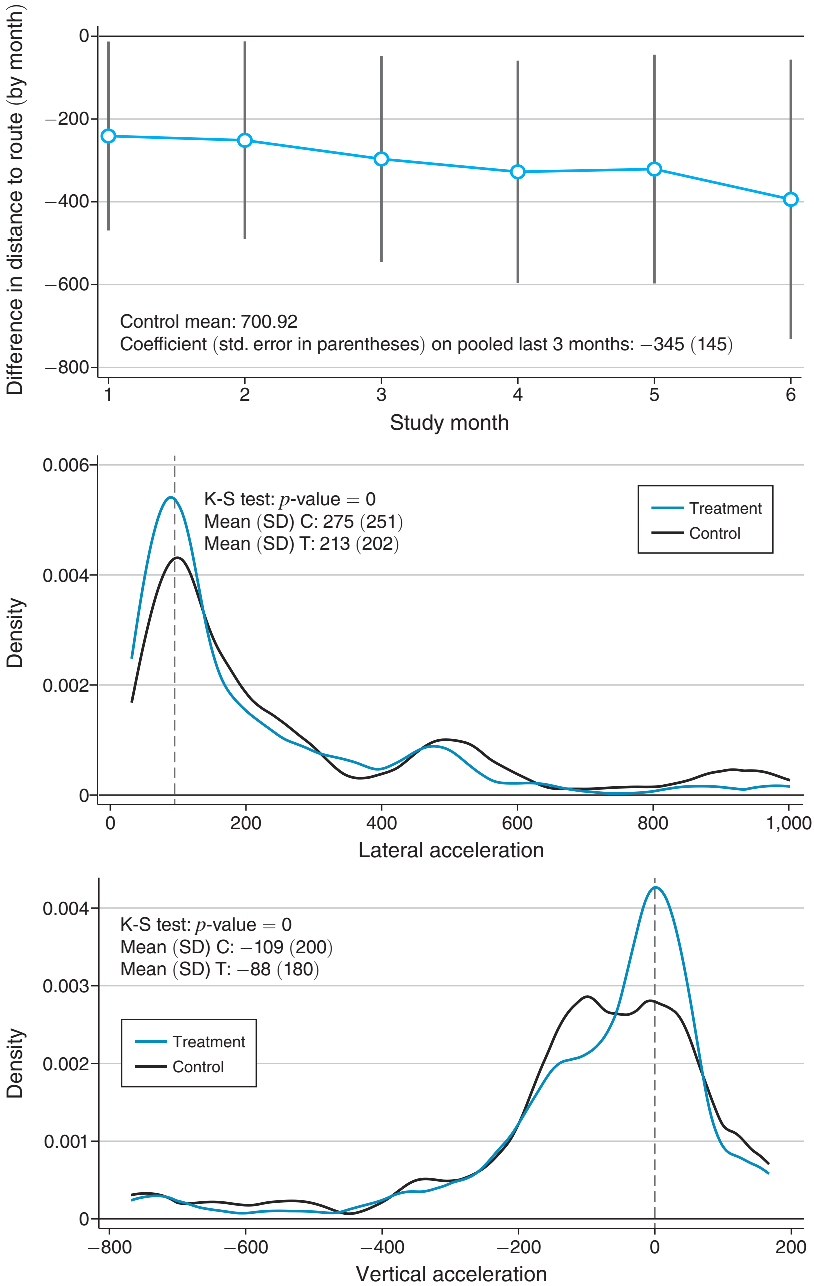

FIGURE 5. TREATMENT EFFECTS ON RISK TAKING

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 percent 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)). Middle panel: Treatment (blue) and control (black) distributions of lateral (left) acceleration. Bottom panel: Treatment (blue) and control (black) distributions of vertical (right) acceleration. We present Kolmogorov-Smirnov tests of equality of these distributions across treatment and control.

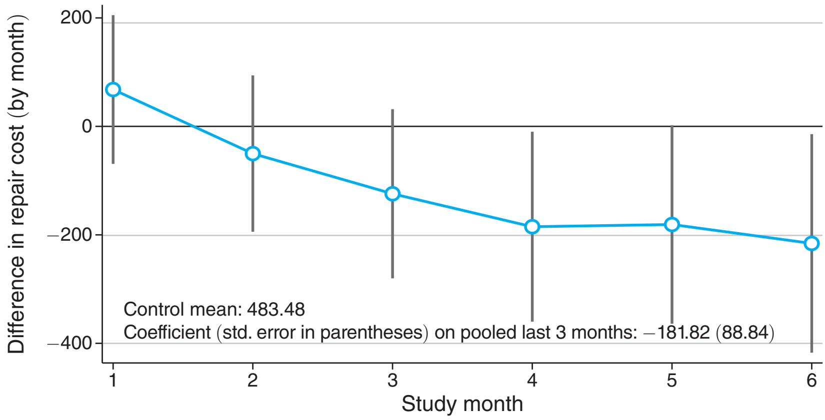

Figure 6. Treatment Effects on Costs

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 percent 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)).

employees' contracts. We also see that drivers are more likely to make the target on occasion (online Appendix Figure A.8 and Table 3, column 7). Drivers are 7 percent more likely to make the target in the third month (p-value $=$ 0.142), 13 percent more likely in the fourth month (p-value $=$ 0.014), and 5 percent more likely in month 6 (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 up front.

Firm Outcomes (*Profits and Growth*).—We now turn to investigating the impact of the monitoring device on firm performance. Firm profits are measured by subtracting costs (repairs and driver salary) from total revenue. Figure 8 (Table 3, column 8) shows that daily profits rise by approximately 13 percent in months 4 and 5—KSh 449 ($4.49) and KSh 453 ($4.53) per day, respectively—for treatment owners (p-value $=$ 0.046 and 0.035, respectively). 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 ($208 if we include variable costs), an amount that could be recovered in less than 3 (4) months. This profit measure does not capture any additional gains from having to spend less time and effort monitoring the driver. We discuss this further in Section VII.

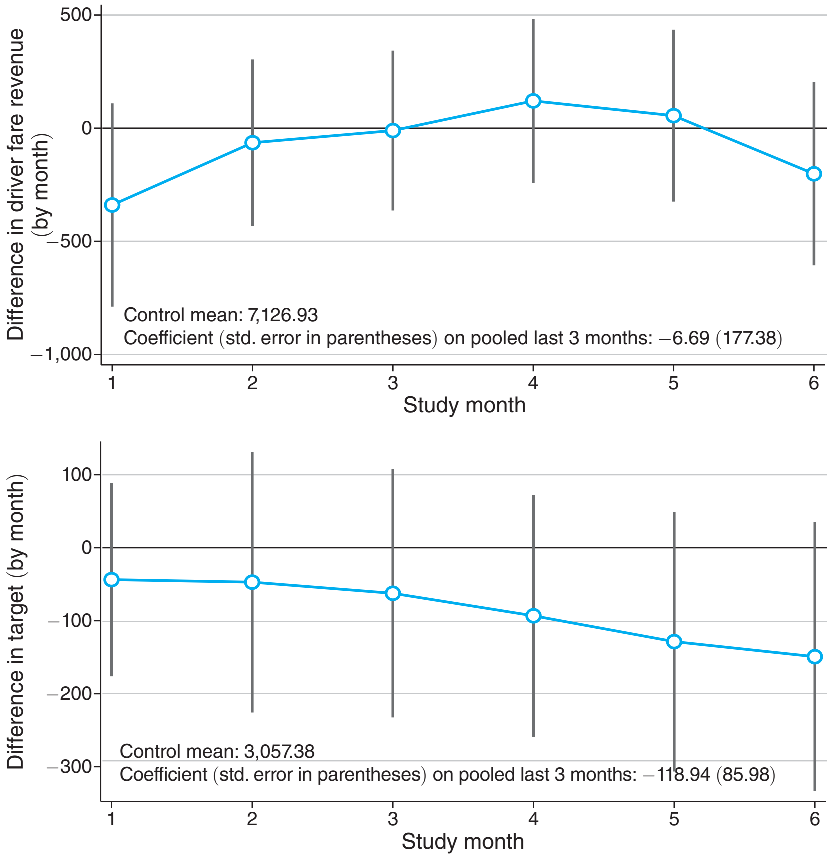

Figure 7. Effect of Monitoring on Revenue and Target

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 percent 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)).

The devices' return on investment suggests they are likely to be a worthwhile investment for owners. 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 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 (Van Lieshout and Mehtha 2017; McKenzie and Woodruff 2017).

Finally, 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 10 percent increase in fleet size (p-value $=$ 0.091). While not statistically significant, Table 4 also demonstrates that owners invested in the interior of their vehicles

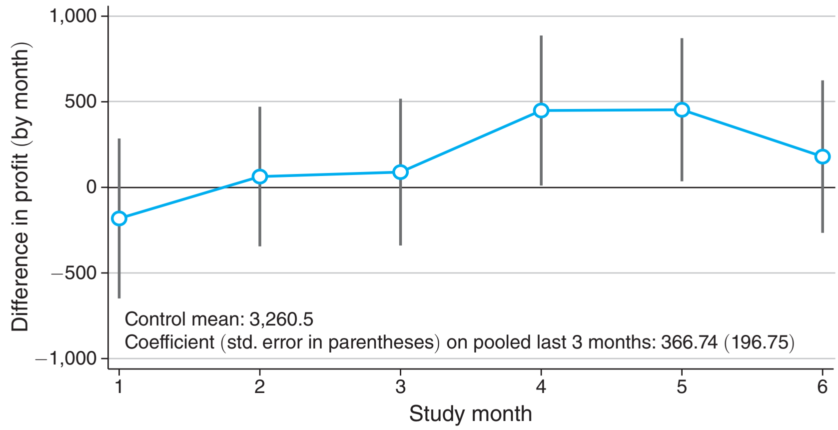

FIGURE 8. TREATMENT EFFECTS ON PROFITS

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 percent 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)).

TABLE 4—TREATMENT EFFECTS ON BUSINESS INVESTMENT

	Number vehicles (1)	New interior (2)
Treatment	0.129 (0.076)	0.074 (0.057)
Control mean of DV	1.22	0.21
Controls	X	X
Route fixed effects	X	X
Matatu observations	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's 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.

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 profit gains that owners reap, which may make it easier to take a loan for a second bus. Second, because the owners likely 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, making it easier to manage. For instance, knowing the number of trips completed by the first matatu can provide

TABLE 5—TREATMENT EFFECTS ON OWNER'S PERCEPTIONS OF THEIR DRIVERS' PERFORMANCE

	Trust amount (1)	More honest (2)	Performance rating (3)	Better driving (4)
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 fixed effects	X	X	X	X
Matatu observations	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's score). "Trust amount": Amount in Kenyan shillings 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 judgment 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.

information about the overall traffic conditions the second matatu faced.[20] 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, columns 4 and 5). Finally, owners' perceptions of their drivers' performance improves by 0.63 points (p-value $= 0.00$) (Table 5). 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.

Driver Outcomes (*Salary and Separation*).—We consider the monetary gains to drivers. Figure 9 shows that the impacts on driver salary per hour are close to zero. The effect of the tracking device on driver take-home pay is ambiguous. The impact depends 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. Finally, we do not see any differences in the rate of separation between drivers and owners in the sample (online 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.[21]

[20] In our 2022 descriptive survey, 50 percent of owners with multiple buses said that having only 1 GPS device in 1 of their minibuses would be sufficient to help them understand the behavior of their other buses.

[21] Note we test whether new drivers are responsible for the 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.

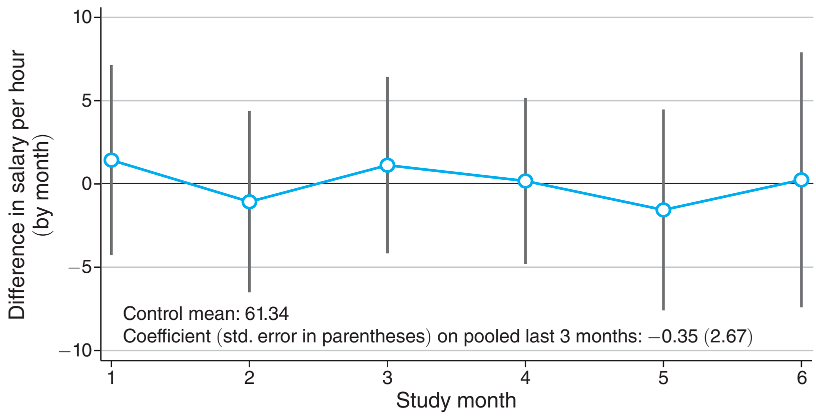

Figure 9. Treatment Effects on Salary per Hour

Notes: OLS estimates according to equation (2). Treatment effect by month on driver salary per hour. Standard errors for 95 percent 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)).

Quantile Treatment Effects.—We explore quantile treatment effects for the core business outcomes (effort, repair costs, revenue, target, profit) at the tenth, twenty-fifth, fiftieth, seventy-fifth, and ninetieth percentile of these distributions (online Appendix 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 ninetieth 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 tenth and twenty-fifth 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 high-revenue days). The median owner experiences these large repairs 12 percent of days their vehicle is on the road, and 90 percent 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. Finally, we see that the impact on the target is largest and statistically significant at the tenth percentile of the distribution, which means some owners are lowering their target below industry norms, rather than owners with high targets readjusting.

VI. 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 IV, 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 affects the welfare of owners and drivers.

A. *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, where 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. The contract characteristics include the rehiring rate, the daily target, and the driver's valuation of the contract. 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. Details about the estimation are presented in online Appendix 6.

Identification.—At baseline (without monitoring), there are three parameters in the model: the driver's outside option, $\bar{u}$; his disutility from his chosen effort and risk, $\psi_B = \psi(e_B, r_B)$; 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 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

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

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

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

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

Its empirical equivalent in the data, that is, the driver's stated value of the contract at baseline, is denoted by U_i.

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

$$T = \underset{\tilde{T} \geq 0}{\arg\max}\, V\big(\tilde{T}\big)$$

with

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

Its empirical equivalent in the data, that is, the observed target, is denoted by T_i. The three structural expressions follow directly from the model after imposing the optimal transfer and rehiring schedules.

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, the 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 online Appendix (online Appendix Tables A.11 to A.18).

Estimation.—We estimate our parameters of interest via generalized method of moments, 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 $\mathbf{X}_i$ consist 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

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

is the vector of moments; T is the optimal target according to the model and the production environment (i.e., revenue and costs), in contrast to the observed target T_i; 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 $\mathbb{E}[p(t)|e_M, r_M]$ 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 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.[22] In both estimations (baseline and monitoring), we can evaluate how well the model-predicted structural components can accommodate the empirical components for both the

[22] 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 to pin down the calibrated parameters.

targeted 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.

B. *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 of 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.[23]

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.[24] Overall, the model does well in matching both the targeted and untargeted moments, which 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 percent 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 desirable.

[23] Online 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 unmatched moments (driver salary, owner profits, and owner value) change as the three parameter estimates $\bar{u}$, $\psi(e,r)$, and h vary.

[24] These moments are linearly dependent, so the gap between observed and predicted salary is always the same magnitude as the gap between observed and predicted profits. One unmodeled 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 the model, with little gain in economic insight.

TABLE 6—MODEL ESTIMATION UNDER BASELINE CONTRACT

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

Input	Value	Interpretation
Panel B. GMM parameter estimates		
Baseline driver disutility $\psi(e_B, r_B)$	2.46	Driving disutility of \$2.46
	(0.75)	
Firing cost h	263	Lost profit of firing of \$263
	(20)	(about 11 days of profit)
Driver outside option $\bar{u}$	1.57	Similar to unskilled daily wage with
	(0.41)	subsistence (\$1.57 + \$7 = \$8.57)

Control group outcome	Reduced-form	Structural	Difference
Panel C. Reduced-form, structural, and matched moments			
Targeted moments			
Rehiring probability $E[p \mid e, r]$	0.007	0.007	−0.000
	(0.000)	(0.001)	(0.001)
Driver contract value U	506.4	507.1	−0.7
	(12.9)	(74.7)	(75.8)
Target T	30.1	30.1	0.0
	(0.4)	(0.8)	(0.9)
Untargeted moments			
Driver salary $E[y \mid e, r] - E[t \mid e, r]$	9.1	9.9	−0.8
	(0.2)	(0.6)	(0.6)
Owner profit $E[t \mid e, r] - E[c \mid e, r]$	24.3	23.5	0.8
	(0.7)	(0.6)	(0.9)
Owner contract value V		2,177	
		(11)	
Welfare $U + V$		2,684	
		(75)	

Notes: GMM estimation of driver disutility, firing cost, and outside option. Sample: Control group. Targeted moments: Rehiring probability, Driver contract value, and Target. "Reduced-form" as observed in the sample. "Structural" is 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.

C. *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 parameter to be estimated is driver disutility under monitoring, $\psi(e_M, r_M)$. In Table 7, panel B, we show its estimated value from the GMM procedure is \$3.74 (SE \$1.74), which is a \$1.27 or 52 percent increase from the disutility estimated under the status quo.

In panel C, we report the observed changes to moments estimated 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

Table 7—Reduced-Form versus Structural Treatment Estimation

Input	Value	Notes	
Panel A. Assumptions			
Subsistence income w	7	Kink in transfer schedule	
Revenue distribution $G(\cdot \mid e, r)$	—	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	
Input	Value	Interpretation	
Panel B. GMM parameter estimates			
Disutility with monitoring $\psi(e_M, r_M)$	3.75	Increase of \$1.28 (52%)	
	(2.61)		
Treatment effect	Reduced-form (Δ)	Structural (Δ)	Difference (Δ)
Panel C. Reduced-form, structural, and matched treatment effects			
Targeted moments			
Rehiring probability $E[p \mid e, r]$	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)
Target T	−1.0	−1.2	0.2
	(0.6)	(1.3)	(1.4)
Untargeted moments			
Driver salary $E[y \mid e, r] - E[t \mid e, r]$	0.1	0.9	−0.9
	(0.2)	(0.9)	(0.9)
Owner profit $E[t \mid e, r] - E[c \mid e, r]$	1.7	0.8	0.9
	(1.0)	(1.4)	(1.7)
Driver contract value U		−19.7	
		(13.8)	
Owner contract value V		82.4	
		(133.5)	
Welfare $U + V$		62.7	
		(137.1)	

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

targeted moments, the structurally estimated treatment effects closely match the observed changes in the firing probability and target. There is a negligible difference in the rehiring probability and a \$1.20 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 statistically 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.[25]

[25] One explanation for this underestimation of owner benefits is that the monitoring device may allow owners to dissuade (unmodeled) driver behavior where they do not give the full revenue amount when it falls below the target

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. 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.

Figure 10 illustrates these dynamics for different costs. When driver disutility increases by less than 20 percent, 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 $-\$19.7$ (SE \$13.8), consistent with the 52 percent estimated increase in $\psi(e,r)$. Meanwhile, we estimate that owner contract value increases by \$82 (SE \$133), leading to a total welfare increase of \$63 (SE \$137). 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.[26]

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.[27]

VII. Discussion of Welfare

While we find that owners offer a welfare-reducing contract to drivers, this did not have to be the case, as we discuss above. Moreover, these welfare estimates do not account for changes in intangible aspects of the relationship between owners and drivers. At baseline, the relationship between owners and drivers is characterized by mutual mistrust. Owners suspect drivers of cheating and reckless driving, while drivers feel that owners second-guess their reports. In focus group discussions, drivers noted that monitoring could improve trust between owners and drivers.

(as discussed 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.

[26] Note that the reduced-form estimate for the effect of monitoring on profits in Section V (\$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.

[27] The willingness-to-pay (WTP) measure is designed to capture owners' present discounted value of profits—though it is worth noting that the procedure used to elicit WTP requires owners to have cash on hand to pay for the devices, and this could bias WTP downward if liquidity constraints limit the amount of money individuals have available (McKenzie and Ubfal 2020).

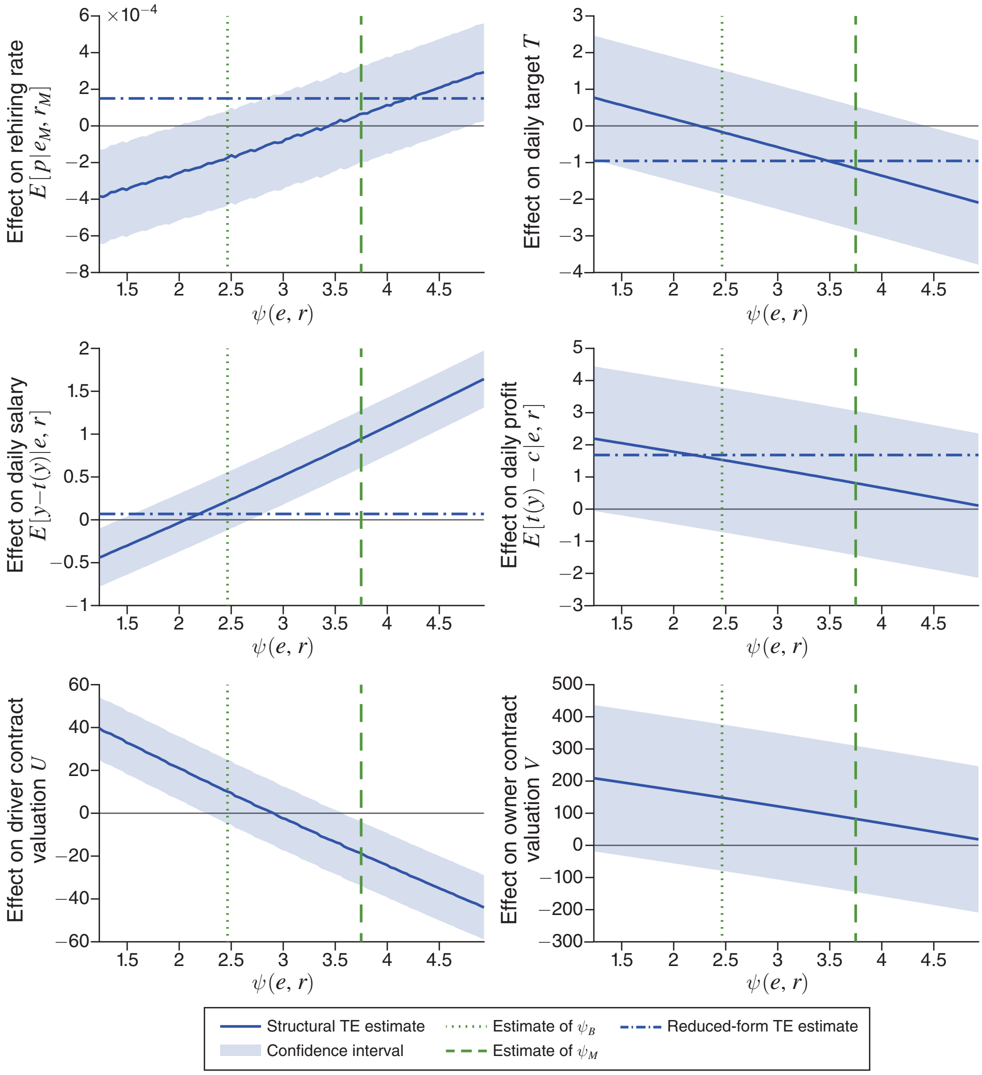

Figure 10. Structural Treatment Effect Estimates Due to Increase in Driver Disutility

Notes: The figure plots the structural treatment effect estimates for driver disutility $\psi(e,r)$ ranging from 50 percent to 200 percent 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. The dotted line shows estimated baseline disutility ψ_B, and the 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.

We complement drivers' welfare estimates with SMS survey responses that we collected from drivers six months after the study finished, when both treatment and control owners had access to the tracking information. Out of the 60 percent of drivers who responded (balanced across treatment and control), one-quarter said the tracking device improved their relationship with the owner, while nearly three-quarters reported no change (only 3 percent reported a worse relationship). Ninety-six percent 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 nonpecuniary 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, 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.

The welfare implications associated with these monitoring technologies are further complicated by how they interact with the public transit passengers and other road users. While beyond the scope of this 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.

VIII. Conclusion

In this paper, we investigate the impact of monitoring devices among small businesses in Kenya. We implement a randomized control trial in which we introduce a monitoring device to 255 firms operating in Nairobi's transit industry. We design a novel mobile application that provides information to 126 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 percent of owners consult the app weekly. Owners also report that monitoring their drivers has become significantly easier.

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 percent, and firm profits increase by 13 percent. 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 percent) 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 remain the norm. This suggests that this class of inefficient contracts could remain widespread in industries where revenue is unobserved, at least until monitoring technologies can reveal the amount of revenue employees earn throughout the day.

We identify the distributional consequences of these technologies by estimating the target contract model via GMM. 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. These losses may be compensated by greater trust between owners and drivers.

Taken together, these results suggest 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 for 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 a 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* 111 (1): 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* 52 (4): 873–87.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro. 2017. "Measuring the Sensitivity of Parameter Estimates to Estimation Moments." *Quarterly Journal of Economics* 132 (4): 1553–92.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2017. "Exporting and Firm Performance: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 132 (2): 551–615.

Baker, George P., and Thomas N. Hubbard. 2004. "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking." *Quarterly Journal of Economics* 119 (4): 1443–79.

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." NBER Working Paper 26993.

Bassi, Vittorio, Raffaela Muoio, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2022. "Achieving Scale Collectively." *Econometrica* 90 (6): 2937–78.

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

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–42.

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* 61 (4): 707–22.

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

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.

Business Daily Africa. 2009. "Vehicle Tracking System Gains Popularity in Kenya," September 16. https://www.businessdailyafrica.com/bd/corporate/companies/vehicle-tracking-system-gains-popularity-in-kenya-1947230.

Caria, Stefano A., and Paolo Falco. Forthcoming. "Skeptical Employers: Experimental Evidence on Biased Beliefs Constraining Firm Growth." *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01219.

Cervero, Robert, and Aaron Golub. 2007. "Informal Transport: A Global Perspective." *Transport Policy* 14 (6): 445–57.

de Rochambeau, Golvine. 2020. "Monitoring and Intrinsic Motivation: Evidence from Liberia's Trucking Firms." Unpublished.

Dickens, William T., Lawrence F. Katz, Kevin Lang, and Lawrence H. Summers. 1989. "Employee Crime and the Monitoring Puzzle." *Journal of Labor Economics* 7 (3): 331–47.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. "Keeping It Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics* 6 (2): 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* 103 (6): 1176–1209.

Gertler, Paul, Marcelo Olivares, Raimundo Undurraga, and Pablo Celhay. 2023. "Monitoring, Organizational Culture, and Procurement Efficiency." Unpublished.

Ghatak, Maitreesh, and Priyanka Pandey. 2000. "Contract Choice in Agriculture with Joint Moral Hazard in Effort and Risk." *Journal of Development Economics* 63 (2): 303–26.

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*: 2–5.

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* 128 (4): 1195–1233.

Habyarimana, James, and William Jack. 2011. "Heckle and Chide: Results of a Randomized Road Safety Intervention in Kenya." *Journal of Public Economics* 95 (11-12): 1438–46.

Habyarimana, James, and William Jack. 2015. "Results of a Large-Scale Randomized Behavior Change Intervention on Road Safety in Kenya." *Proceedings of the National Academy of Sciences* 112 (34): E4661–70.

Hanna, Rema, Gabriel Kreindler, and Benjamin A. Olken. 2017. "Citywide Effects of High-Occupancy Vehicle Restrictions: Evidence from "three-in-one" in Jakarta." *Science* 357 (6346): 89–93.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization* 7: 24.

Hsieh, Chang-Tai, and Benjamin A. Olken. 2014. "The Missing "Missing Middle"." *Journal of Economic Perspectives* 28 (3): 89–108.

Hubbard, Thomas N. 2000. "The Demand for Monitoring Technologies: The Case of Trucking." *Quarterly Journal of Economics* 115 (2): 533–60.

Hubbard, Thomas N. 2003. "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking." *American Economic Review* 93 (4): 1328–53.

Innes, Robert D. 1990. "Limited Liability and Incentive Contracting with Ex-Ante Action Choices." *Journal of Economic Theory* 52 (1): 45–67.

Jayachandran, Seema. 2020. "Microentrepreneurship in Developing Countries." Unpublished.

Kalan, Jonathan. 2013. "The Technology Modernising Kenya's Matatus." *BBC,* October 8. https://www.bbc.com/news/business-24426676.

Karlan, Dean, and Martin Valdivia. 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *Review of Economics and Statistics* 93 (2): 510–27.

Kelley, Erin M., Gregory Lane, and David Schönholzer. 2024. *Data and Code for: "Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E202063V1.

Kenya Roads Board. 2007. *Kenyan Transport Sector Details Annex 3.1*. Nairobi, Kenya: Kenya Roads Board.

Kreindler, Gabriel. 2020. "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications." Unpublished.

Lane, Gregory, David Schönholzer, and Erin M. Kelley. 2022. "Information and Strategy in Lemon Markets: Improving Safety in Informal Transit." Unpublished.

Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review* 93 (3): 835–57.

Liu, Meng, Erik Brynjolfsson, and Jason Dowlatabadi. 2021. "Do Digital Platforms Reduce Moral Hazard? The Case of Uber and Taxis." *Management Science* 67 (8): 4665–85.

Lucas, Jr., Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2): 508–23.

Macharia, W. M., E. K. Njeru, F. Muli-Musiime, and V. Nantulya. 2009. "Severe Road Traffic Injuries in Kenya, Quality of Care and Access." *African Health Sciences* 9 (2): 118–24.

McCormick, Dorothy, Winnie Mitullah, Preston Chitere, Risper Orero, and Marilyn Om-meh. 2013. "Paratransit Business Strategies: A Bird's-Eye View of Matatus in Nairobi." *Journal of Public Transportation* 16 (2): 135–52.

McKenzie, David, and Anna Luisa Paffhausen. 2019. "Small Firm Death in Developing Countries." *Review of Economics and Statistics* 101 (4): 645–57.

McKenzie, David, and Diego Ubfal. 2020. "Using BDM and TIOLI to Measure the Demand for Business Training in Jamaica." Unpublished.

McKenzie, David, and Christopher Woodruff. 2017. "Business Practices in Small Firms in Developing Countries." *Management Science* 63 (9): 2967–81.

Mutongi, Kenda. 2017. *Matatu: A History of Popular Transportation in Nairobi.* Chicago, IL: 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* 10 (1-2): 53–61.

Shahe Emran, M., A. K. M. Mahbub Morshed, and Joseph E. Stiglitz. 2021. "Microfinance and Missing Markets." *Canadian Journal of Economics* 54 (1): 34–67.

Sprenger, Charles. 2015. "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points." *Journal of Political Economy* 123 (6): 1456–99.

Startz, Meredith, Gregory Fischer, Dean Karlan, and Diego Santa Maria. 2023. "Business Training and Spatial Competition in Urban Uganda." Unpublished.

Townsend, Robert M. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (2): 265–93.

Tsivanidis, Nick. 2019. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio." Unpublished.

US Department of Transportation. 2016. *National Transportation Statistics*. Washington, DC: Bureau of Transportation Statistics.

Van Lieshout, Susanne, and Pranati Mehtha. 2017. *The Next 15 Million: Start and Improve Your Business Global Tracer Study 2011-2015*. Geneva: International Labour Organization.

West, Darell M. 2021. "How Employers use Technology to Surveil Employees." *Brookings Techtank,* January 5. https://www.brookings.edu/articles/how-employers-use-technology-to-surveil-employees/.

WHO. 2013. *Global Status Report on Road Safety 2013: Supporting a Decade of Action*. Geneva, Switzerland: 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*. Washington, DC: World Bank.